Regularization of Nonparametric Frontier Estimators

Abdelaati Daouia[1,2]
daouia@cict.fr

Jean-Pierre Florens[1]
florens@cict.fr

Léopold Simar[1,2]*
leopold.simar@uclouvain.be

[1] Toulouse School of Economics, France
[2] Institut de Statistique, Université catholique de Louvain, Belgium

Abstract

In production theory and efficiency analysis, we estimate the production frontier, the locus of the maximal attainable level of an output (the production), given a set of inputs (the production factors). In other setups, we estimate rather an input (or cost) frontier, the minimal level of the input (cost) attainable for a given set of outputs (goods or services produced). In both cases the problem can be viewed as estimating a surface under shape constraints (monotonicity, ...). In this paper we derive the theory of an estimator of the frontier having an asymptotic normal distribution. It is based on the order-m partial frontier where we let the order m to converge to infinity when $n \to \infty$ but at a slow rate. The final estimator is then corrected for its inherent bias. We thus can view our estimator as a regularized frontier. In addition, the estimator is more robust to extreme values and outliers than the usual nonparametric frontier estimators, like FDH and than the unregularized order-m_n estimator of Cazals et al. (2002) converging to the frontier with a Weibull distribution if $m_n \to \infty$ fast enough when $n \to \infty$. The performances of our estimators are evaluated in finite samples and compared to other estimators through some Monte-Carlo experiments, showing a better behavior (in terms of robustness, bias, MSE and achieved coverage of the resulting confidence intervals). The practical implementation and the robustness properties are illustrated through simulated data sets but also with a real data set.

Key words : Production function, Free Disposal Hull, Nonparametric frontier, Robust estimation, Extreme value index.

*This research was supported by the French "Agence Nationale pour la Recherche" under grant ANR-08-BLAN-0106-01/EPI project. Support from the "Interuniversity Attraction Pole", Phase VI (No. P6/03) of the Belgian Science Policy and from the Chair of Excellency "Pierre de Fermat", Région Midi-Pyrénées, France are also acknowledged.

1 Introduction

In production theory and efficiency analysis, we are interested in estimating the production frontier which is the locus of the maximal attainable level of an output (the production), given a set of inputs (the production factors). In other setups, we are rather willing to estimate an input (or cost) frontier that is defined as the minimal attainable level of the input (cost) for a given set of outputs (goods or services produced). In both cases the problem can be viewed as estimating a surface under shape constraints (monotonicity,...). The efficiency score of a given unit is then determined by an appropriate distance (in the output direction, or in the input direction) of this unit to the optimal frontier.

Formally (we will in this paper focus the presentation in the input orientation case, where we want to estimate the minimal cost frontier[1]), let $x \in \mathbb{R}_+$ denote the input (or the cost of production) and $y \in \mathbb{R}_+^q$ be the vector of goods or services produced. The attainable set (feasible combinations of input and outputs) is defined as

$$\Psi = \{(x, y) \in \mathbb{R}_+ \times \mathbb{R}_+^q \mid y \text{ can be produced by } x\}. \tag{1.1}$$

A minimal assumption often accepted for Ψ is the free disposability of the inputs and of the outputs, namely, if $(x, y) \in \Psi$, then $(x', y') \in \Psi$ for any pairs (x', y') such that $x' \geq x$ and $y' \leq y$. This implies a monotonicity property of the frontier surface. Sometimes (not in this paper), the hypothesis of the convexity of Ψ is also assumed (see Shephard, 1970 for a comprehensive overview of the underlying economic models used in prodution theory). The efficient boundary of Ψ, in the input oriented case, is represented by the minimal input frontier function

$$\varphi(y) = \inf\{x \mid (x, y) \in \Psi\}, \tag{1.2}$$

and the Farrell-Debreu efficiency score of a unit operating at the level (x_0, y_0) is given by the ratio $\varphi(y_0)/x_0$, which gives a number between zero and one. An efficiency equal to one corresponds to an input-efficient unit (being on the minimal input frontier) and more generally $\varphi(y_0)/x_0 \leq 1$ gives the reduction of input (cost) the firm should reach to be considered as input-efficient.

A popular nonparametric estimator of the attainable set is the Free Disposal Hull (FDH) estimator proposed by Deprins, Simar and Tulkens (1984). The FDH is the smallest monotone set enveloping the data points, it relies only on the free disposability assumption and its asymptotic properties have been established (Park, Simar and Weiner, 2000 and Daouia,

[1]The presentation for the output oriented case, where we want to estimate the maximal production frontier, is a straightforward adaptation of what is done here. In the appendix, we give a summary of the notations and main results for that case.

Florens, Simar, 2010). More details will be given below. Another nonparametric estimator, the Data Envelopment Analysis (DEA), initiated by Farrell (1957) and popularized by Charnes, Cooper and Rhodes (1978), can be justified when the convexity of Ψ is moreover assumed. Its asymptotic properties have been established in Kneip et al. (2008). A recent survey of the available statistical tools for making inference in these nonparametric models can be found in Simar and Wilson (2008).

In most of the empirical examples, a naive application of these nonparametric techniques is impossible because real samples contain in general some anomalous data. In that case, the estimated frontier is fully determined by these outliers and the measurement of inefficiencies are totally unrealistic. Whereas most of the practitioners use a rule of thumb for outliers elimination, we have rather proposed in previous different papers (Cazals et al., 2002 and Daouia and Simar, 2007) to keep all the observations in the sample but to replace the frontier of the empirical distribution by (conditional) quantiles or by the expectation of the minimum (or maximum) of a subsample of the data. This latter method defines the order-m frontier. The underlying idea of the two existing methods is thus to estimate a partial frontier well inside the cloud of data points but near its lower (or upper) boundary, in such a way to be sensitive to the magnitude of the extreme valuable observations but, simultaneously, resistant to their influence in case they are suspicious (see Daouia and Gijbels, 2011a, for additional justifications). This is an efficient estimation strategy in absence of information on whether the (isolated) observations are measured accurately. The duality between order-m and order-α frontiers has been also investigated by Daouia and Gijbels (2011a). They show in particular, that even if the order-α quantile frontiers have global better robustness properties (higher breakdown value), it appears that once they breakdown, they become less resistant to outliers than the order-m frontiers.

It has been proved in Cazals et al. (2002) that if m is fixed the nonparametric estimator of the order-m frontier is asymptotically normal with a parametric $\sqrt{n}$ rate of convergence. However, if m goes to infinity very fast as n tends to infinity, the asymptotic behavior of the resulting estimator is identical to the asymptotics of the FDH estimator (Weibull limiting distribution, with the "curse of the dimensionality"). On the other hand, even in absence of outliers, the order-m frontier estimator is biased. The methodology we propose is to increase slowly m when n increases in order to keep the asymptotic normality (as the case of fixed m) and then to correct the bias. The advantage of this approach is to be unbiased in absence of outliers (which is not the case of the FDH estimator) and to be less sensitive in case of anomalous data in the sample.

The outliers may come from measurement errors, from reporting errors in the survey, or from unobserved heterogeneity in the sample. One may imagine that the fraction of outlying

observations comes from a totally different process of the main data generating process, and the objective is then to estimate the frontier of the main production process.

Our method requires the estimation of the tail index of the distribution (defined and denoted ρ_y below) and a coefficient ℓ_y (defined below) which describe the behavior of the distribution close to its support frontier. We can view (ρ_y, ℓ_y) as regularity parameters. These estimations and the estimation of the frontier itself need the selection of "regularization" parameters (m depending on n and a particular sample fraction a). This selection (of the order m in particular) naturally implies a tradeoff between the bias of the estimator (even if it is corrected at the first order) and the sensitivity to the outliers. To the best of our knowledge, this specification of regularity parameters balancing bias and robustness in a decisional framework has not yet been proposed and it is in our research agenda from a perspective of extreme-value theory.

Our approach is exactly in the spirit of nonparametric estimation: we estimate a function under some regularity conditions on the underlying distribution. These conditions depend on the parameters (ρ_y, ℓ_y). If these parameters are known, the nonparametric estimation is easy, but in general such parameters should be preliminarily estimated. We propose here a method that appears to be more efficient in the simulation experiments than previous ones suggested in the literature.

The other concept of partial frontier mentionned above is the order-α (conditional) quantile frontier (see Aragon et al. 2005 and Daouia and Simar, 2007), providing an alternative robust estimator of the frontier function. Recent work of Daouia et al. (2010) makes the links between frontier estimation and extreme-value theory. By doing so, they revisit and extend former results on the asymptotic behavior of the FDH estimator and they provide an alternative regularized version of the FDH estimator by using extreme quantile-type frontiers (see also the footnote 3 below). Here too, an estimation of the regularity parameters (ρ_y, ℓ_y) is proposed showing however a greater unstability with respect to the choice of the regularization parameter $\alpha = \alpha_n$. The quantile-based method could also require huge sample sizes to get reasonable estimates of the tail index. We will show in the simulation exercises that the new method developped in the paper at hand provides much better estimates.

The next section gives the basic notation for introducing the FDH and order-m frontier estimators. Section 3 provides the main theoretical results of this paper: the estimation of the order-m frontier when m tends to infinity (subsection 3.1) and how to implement an estimator of the frontier function $\varphi(y)$ in practice (subsection 3.2). Section 4 addresses the problem of estimating the unknown parameters of the asymptotic distribution. Section 5 illustrates how our procedure works out in practice with simulated data and a real data set. Section 6 concludes.

2 Basic Concepts

2.1 The FDH estimator

The attainable set Ψ can be seen as the support of the random vector (X, Y) defined on an appropriate probability space. It will be useful to describe the joint distribution of (X, Y) by its joint survivor function:

$$S_{XY}(x, y) = \text{Prob}(X \geq x, Y \geq y) = S(x|y)S_Y(y), \tag{2.1}$$

where $S(x|y) = \text{Prob}(X \geq x \,|\, Y \geq y)$ and $S_Y(y) = \text{Prob}(Y \geq y)$. Notice that the conditional survivor function $S(x|y)$ is non-standard, since the condition is $Y \geq y$.

Cazals, Florens and Simar (2002) have shown that under the free disposability assumption, the minimal input function $\varphi(y)$ can equivalently be defined as

$$\varphi(y) = \inf\{x \,|\, S(x|y) < 1\}. \tag{2.2}$$

Since the attainable set is unknown, it has to be estimated from a sample of i.i.d. units $\mathcal{X}_n = \{(X_i, Y_i) \,|\, i = 1, \ldots, n\}$. The free disposal hull of $\mathcal{X}_n$ is the FDH estimator

$$\widehat{\Psi} = \{(x, y) \mid y \leq Y_i, x \geq X_i, i = 1, \ldots, n\}, \tag{2.3}$$

providing the FDH estimator of the frontier $\varphi(y)$

$$\hat{\varphi}(y) = \inf\{x \,|\, \widehat{S}(x|y) < 1\} = \min_{\{i:Y_i \geq y\}} X_i, \tag{2.4}$$

where $\widehat{S}(x|y) = \widehat{S}_{XY}(x, y)/\widehat{S}_Y(y)$ with $\widehat{S}_{XY}(x, y) = (1/n)\sum_{i=1}^n \mathbb{1}(X_i \geq x, Y_i \geq y)$ and $\widehat{S}_Y(y) = (1/n)\sum_{i=1}^n \mathbb{1}(Y_i \geq y)$. Park et al. (2000) have obtained the limiting distribution of FDH estimators in a full multivariate set-up under some regularity conditions. The most general asymptotic result in our setup here is given by Daouia et al. (2010) and can be summarized as follows.

Under the regularity condition (Corollary 2.2 in Daouia et al., 2010)

$$S_Y(y)(1 - S(x|y)) = \ell_y\big(x - \varphi(y)\big)^{\rho_y} + o\big((x - \varphi(y))^{\rho_y}\big), \text{ as } x \downarrow \varphi(y), \tag{2.5}$$

with $\ell_y > 0$, $\rho_y > q$ and $\varphi(y)$ being differentiable in y with strictly positive first partial derivatives, we have[2] as $n \to \infty$

$$(n\ell_y)^{1/\rho_y}\big(\hat{\varphi}(y) - \varphi(y)\big) \xrightarrow{\mathcal{L}} \text{Weibull}(1, \rho_y). \tag{2.6}$$

[2]The Weibull distribution is related to the Exponential distribution: $W \sim \text{Weibull}(1, c) \Leftrightarrow W^c \sim \text{Exp}(1)$.

In addition, the joint density of (X, Y) near the frontier function can be expressed as

$$f(x, y) = c_y(x - \varphi(y))^{\beta_y} + o\big((x - \varphi(y))^{\beta_y}\big), \text{ as } x \downarrow \varphi(y), \tag{2.7}$$

where $c_y > 0$ and $\beta_y = \rho_y - (q + 1)$. Since $\beta_y > -1$, the asymptotic result covers the cases $-1 < \beta_y < 0$, where the density tends to infinity at the frontier, at a speed of the power $\rho_y - (q + 1)$, the case $\beta_y = 0$ where the density has a jump at the frontier ($\rho_y = q + 1$) and the cases $\beta_y > 0$ where the joint density decays to zero at a speed of the power $\rho_y - (q + 1)$.

Remark 2.1. *The regularity condition* (2.5) *is a particular case of the more general extreme value regularity condition (see Daouia et al., 2010 for details)*

$$S_Y(y)(1 - S(x|y)) = L_y\left(\frac{1}{x - \varphi(y)}\right)(x - \varphi(y))^{\rho_y}, \tag{2.8}$$

where L_y is a slowly varying function and $\rho_y > 0$ is the tail index, i.e. $\lim_{t\to\infty} \frac{L_y(tz)}{L_y(t)} = 1$, for all $z > 0$. This is the necessary and sufficient condition under which the conventional FDH estimator converges to a non-degenerate distribution.

Given that the regularly varying function L_y in (2.8) *satisfies $L_y(tz) \approx L_y(t)$ as $t \to \infty$, for all $z > 0$, a more convenient condition is to consider the normalized class of slowly varying functions $L_y(z)$ that can be approximated by a constant $\ell_y > 0$ for all z large enough, or equivalently $L_y\left(\frac{1}{x-\varphi(y)}\right) \approx \ell_y$ as $x \downarrow \varphi(y)$. This will be sufficient to ensure the asymptotic Normality of our regularized version of the FDH estimator. In this case, the necessary and sufficient condition entails to equation* (2.7) *which is a very common assumption in the statistical literature on frontier estimation (see Daouia et al., 2010 and the references therein). In the econometric literature on frontier analysis, the shape parameter β_y of the joint density is simply set equal to zero in most of the nonparametric approaches, while (β_y, ℓ_y) are assumed to be known in all the parametric approaches.*

For instance, if (X, Y) is uniformly distributed over $\Psi = \{(x, y)|0 \le y \le x \le 1\}$, we have $L_y(\cdot) = \ell_y = \ell = 1$ and $\rho_y = \rho = 2$ and (2.5) *is satisfied.*

If $X = Y^{1/2}\exp(U)$ where Y is uniform over $[0, 1]$ and U, independent of Y, is Exponential with parameter $\lambda = 3$, we have $\rho_y = \rho = 2$ and $L_y\left(\frac{1}{x-\varphi(y)}\right) = \ell_y + o\big((x - \varphi(y))\big)$ when $x \downarrow \varphi(y)$, with $\ell_y = \ell = 3$ and (2.5) *is satisfied.*

2.2 Order-m frontier and robust estimation

By construction, since it envelops all the data points, the FDH estimator (and its convexified version, the DEA estimator) is very sensitive to outliers and extreme data points. Cazals et al. (2002) suggested to define a benchmark frontier that is less extreme than the full

frontier function $\varphi(y)$. Indeed, the latter can be defined as the minimal achievable input level for firms producing at least the level y, see (2.2). A less extreme benchmark, based on the concept of order-m frontier, is defined as the expected minimal input value among m peers drawn at random in the population of units producing at least the level y, where m is a natural number ($m \geq 1$)[3]. Formally,

$$\varphi_m(y) = \mathbb{E}\left[\min(X_1, \ldots, X_m) | Y \geq y\right], \tag{2.9}$$

provided the expectation exists. We have the following equivalences

$$\varphi_m(y) = \int_0^\infty S^m(u|y)\, du = \varphi(y) + \int_{\varphi(y)}^\infty S^m(u|y)\, du. \tag{2.10}$$

It can be seen that $\varphi_m(y) \to \varphi(y)$ as $m \to \infty$. Recently, Daouia and Gijbels (2011b, Proposition 1) have shown that $\varphi_m(y)$ exists, for all $m \geq 1$, provided that $\mathbb{E}\left[X|Y \geq y\right] < \infty$.

A nonparametric estimator of $\varphi_m(y)$ is given by pluging the empirical version of $S(u|y)$ in (2.10) to obtain

$$\hat{\varphi}_m(y) = \int_0^\infty \widehat{S}^m(u|y)\, du. \tag{2.11}$$

For fixed m, it has been shown that $\sqrt{n}\big(\hat{\varphi}_m(\cdot) - \varphi_m(\cdot)\big) \xrightarrow{\mathcal{L}} \mathcal{G}(0, \Omega)$ where $\mathcal{G}$ is a gaussian process with covariance function Ω described in Cazals et al. (2002). In particular, for any given y and a fixed value of m, we have as $n \to \infty$,

$$\frac{\sqrt{n}}{\sigma(m,y)}\big(\hat{\varphi}_m(y) - \varphi_m(y)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \tag{2.12}$$

where

$$\sigma^2(m,y) = \mathbb{E}\left[\frac{m\, \mathbb{1}(Y \geq y)}{S_Y(y)} \int_0^\infty \Big(S^{m-1}(u|y)\, \mathbb{1}(X \geq u) - S^m(u|y)\Big)\, du\right]^2. \tag{2.13}$$

It is clear that if $m \to \infty$, $\hat{\varphi}_m(y)$ will converge to the FDH estimator $\hat{\varphi}(y)$. Cazals et al. show that if $m = m_n \to \infty$ fast enough when $n \to \infty$, the resulting estimator has the same asymptotic distribution as the FDH estimator, *i.e.*,

$$(n\ell_y)^{1/\rho_y}\big(\hat{\varphi}_{m_n}(y) - \varphi(y)\big) \xrightarrow{\mathcal{L}} \text{Weibull}(1, \rho_y). \tag{2.14}$$

Of course, for finite n, the resulting estimator $\hat{\varphi}_{m_n}(y)$ does not envelop all the data points and so provides a robust version of the FDH estimator.

[3] To help the reader to understand the difference with the quantile-type frontiers of Aragon et al. (2005), and used in Daouia et al. (2010) as an alternative regularization of the FDH estimator, we briefly remind that $\varphi_\alpha(y) = \inf\{x | S_{X|Y}(x|y) < \alpha\}$, where $\alpha \in (0, 1]$. The nonparametric estimator $\widehat{\varphi}_\alpha(y)$ is obtained by pluging $\widehat{S}_{X|Y}$ at the place of $S_{X|Y}$. When $\alpha_n \to 1$ slowly enough as $n \to \infty$, Daouia et al. (2010) derive, after a bias correction, a regularized estimator of $\varphi(y)$ having also a normal limiting distribution.

3 The Main Results

3.1 Estimation of the order-m frontier when $m \to \infty$

We start with a preliminary lemma which controls, as $m \to \infty$, the variance of the order-m estimator $\hat{\varphi}_m(y)$ given in (2.13).

Lemma 3.1. *Under the regularity condition* (2.5)*, we have for any y such that $S_Y(y) > 0$, as $m \to \infty$*

$$k_{1,y}\, m^{1-2/\rho_y} \leq \sigma^2(m,y) \leq k_{2,y}\, m^{2-2/\rho_y}, \tag{3.1}$$

where $k_{1,y}$ and $k_{2,y}$ are some positive constants.

Proof: We first obtain after some elementary algebraic manipulations that the variance can be expressed as

$$\sigma^2(m,y) = \frac{2m^2}{S_Y(y)} \int_{\varphi(y)}^{\infty} \int_{\varphi(y)}^{\infty} S^m(u|y) S^{m-1}(v|y)(1 - S(v|y))\, \mathbb{1}(u \geq v)\, du\, dv. \tag{3.2}$$

(i) Searching a minorant of $\sigma^2(m,y)$ when $m \to \infty$. We first notice that

$$\sigma^2(m,y) = \frac{2m^2}{S_Y(y)} \int_{\varphi(y)}^{\infty} S^{m-1}(v|y) F(v|y) \left[\int_v^{\infty} S^m(u|y)\, du \right] dv,$$

where $F(v|y) = 1 - S(v|y)$. So that for all $\delta > 0$, we have

$$\begin{aligned} \sigma^2(m,y) &\geq \frac{2m^2}{S_Y(y)} \int_{\varphi(y)}^{\varphi(y)+\delta} S^{m-1}(v|y) F(v|y) \left[\int_v^{v+\delta} S^m(u|y)\, du \right] dv, \\ &\geq \frac{2m^2\delta}{S_Y(y)} \int_{\varphi(y)}^{\varphi(y)+\delta} S^{m-1}(v|y) F(v|y) S^m(v+\delta)\, dv. \end{aligned}$$

Since $S^{m-1}(v|y) \geq S^{m-1}(v+\delta|y) \geq S^m(v+\delta|y)$, we have

$$\begin{aligned} \sigma^2(m,y) &\geq \frac{2m^2\delta}{S_Y(y)} \int_{\varphi(y)}^{\varphi(y)+\delta} S^{2m}(v+\delta|y) F(v|y)\, dv, \\ &\geq \frac{2m^2\delta}{S_Y(y)} S^{2m}(\varphi(y)+2\delta|y) \int_{\varphi(y)}^{\varphi(y)+\delta} F(v|y)\, dv. \end{aligned}$$

Now, if $\delta \downarrow 0$, by the regularity condition (2.5) we have that

$$\int_{\varphi(y)}^{\varphi(y)+\delta} F(v|y)\, dv \geq \frac{c_y}{\rho_y + 1} \frac{\delta^{\rho_y+1}}{2}, \tag{3.3}$$

where $c_y = \dfrac{\ell_y}{S_Y(y)}$. When $\delta \downarrow 0$, it is also easy to see from (2.5) that

$$S(\varphi(y)+2\delta|y) \;\geq\; 1-2c_y(2\delta)^{\rho_y} = \exp\Big[\log\big(1-2c_y(2\delta)^{\rho_y}\big)\Big].$$

Therefore $S^{2m}(\varphi(y)+2\delta|y) \geq \exp\Big[2m\log\big(1-2c_y(2\delta)^{\rho_y}\big)\Big]$. Since $\lim\limits_{\delta\downarrow 0}\dfrac{\log\big(1-2c_y(2\delta)^{\rho_y}\big)}{-2c_y(2\delta)^{\rho_y}} = 1$, for sufficiently small $\delta > 0$ we have $\dfrac{\log\big(1-2c_y(2\delta)^{\rho_y}\big)}{-2c_y(2\delta)^{\rho_y}} \leq 2$. So, when $\delta \downarrow 0$ we have $S^{2m}(\varphi(y)+2\delta|y) \geq e^{-8mc_y(2\delta)^{\rho_y}}$. Pluging these results in the latter inequality for $\sigma^2(m,y)$ we have as $\delta \downarrow 0$

$$\sigma^2(m,y) \geq \frac{2m^2\delta}{S_Y(y)}\, e^{-8mc_y(2\delta)^{\rho_y}}\, \frac{c_y}{\rho_y+1}\, \frac{\delta^{\rho_y+1}}{2}.$$

Choosing $\delta = (1/m)^{1/\rho_y}$, we have as $m \to \infty$

$$\sigma^2(m,y) \geq k_{1,y}\, m^{1-2/\rho_y}. \tag{3.4}$$

(ii) Searching a majorant of $\sigma^2(m,y)$ when $m \to \infty$. From (2.10) and (3.2) we have

$$\begin{aligned}
\sigma^2(m,y) &\leq \frac{2m^2}{S_Y(y)} \int_{\varphi(y)}^{\infty}\int_{\varphi(y)}^{\infty} S^m(u|y)S^{m-1}(v|y)(1-S(v|y))\,du\,dv, \\
&= \frac{2m^2}{S_Y(y)}\left[(\varphi_m(y)-\varphi(y))(\varphi_{m-1}(y)-\varphi(y)) - (\varphi_m(y)-\varphi(y))^2\right] \\
&= \frac{2m^2}{S_Y(y)}(\varphi_m(y)-\varphi(y))^2\left[\frac{\varphi_{m-1}(y)-\varphi(y)}{\varphi_m(y)-\varphi(y)} - 1\right]
\end{aligned}$$

Now, by the regularity condition (2.5), the equation (2.5) in Daouia et al. (2010) and from the definition (2.9) of φ_m, we have as $m \to \infty$

$$\varphi_m(y) - \varphi(y) = \Gamma\left(1+\frac{1}{\rho_y}\right)\left(\frac{1}{m\,\ell_y}\right)^{1/\rho_y} + o(m^{-1/\rho_y}). \tag{3.5}$$

Therefore, as $m \to \infty$,

$$\sigma^2(m,y) \;\leq\; \frac{2m^2}{S_Y(y)}\left[\Gamma^2(1+1/\rho_y)(m\ell_y)^{-2/\rho_y} + o(m^{-2/\rho_y})\right] \leq k_{2,y}m^{2-2/\rho_y},$$

where $k_{2,y}$ is a positive constant. This completes the proof of the lemma. □

The following theorem gives the basic results of our paper, it specifies under which condition on the sequence m_n, the asymptotic distribution of $\hat{\varphi}_{m_n}(y)$ is still a Normal distribution.

Theorem 3.1. *Under the regularity condition* (2.5)*, and if* $m_n = c\, n^{1/3-\varepsilon}\,(\log\log n)^{-2/3}$ *for some constants* $c > 0$ *and* $\varepsilon \in (0, 1/3)$*, we have for any* y *such that* $S_Y(y) > 0$*, as* $n \to \infty$

$$\frac{\sqrt{n}}{\sigma(m_n, y)}\big(\hat{\varphi}_{m_n}(y) - \varphi_{m_n}(y)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{3.6}$$

Proof: In the proof, to simplify the notation, we will denote m_n by m, keeping in mind that $m = m_n \to \infty$ when $n \to \infty$ at the rate given by m_n. Let us define

$$R^y_{m,n} = \big(\hat{\varphi}_m(y) - \varphi_m(y)\big) - m \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\big[\widehat{S}(u|y) - S(u|y)\big]\, du.$$

So the object of interest for the theorem can be written as

$$\begin{aligned}\frac{\sqrt{n}}{\sigma(m, y)}\big(\hat{\varphi}_m(y) - \varphi_m(y)\big) &= \frac{m\sqrt{n}}{\sigma(m, y)} \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\big[\widehat{S}(u|y) - S(u|y)\big] du \\ &+ \frac{\sqrt{n}}{\sigma(m, y)} R^y_{m,n}. \end{aligned} \tag{3.7}$$

(i) We first prove that $\frac{\sqrt{n}}{\sigma(m,y)}\, R^y_{m,n} \xrightarrow{a.s.} 0$ *as* $n \to \infty$. Since $\hat{\varphi}(y) \overset{a.s.}{\geq} \varphi(y)$, and because $\widehat{S}^m(u) = S^m(u) = 1$ for all $u \in (0, \varphi(y))$, we have

$$R^y_{m,n} \overset{a.s.}{=} \int_{\varphi(y)}^{\infty} \big(\widehat{S}^m(u|y) - S^m(u|y)\big)\, du - m \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\big[\widehat{S}(u|y) - S(u|y)\big]\, du.$$

Now, consider the following Taylor expansion

$$\begin{aligned}\int_{\varphi(y)}^{\infty} \big(\widehat{S}^m(u|y) - S^m(u|y)\big)\, du &= m \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\big[\widehat{S}(u|y) - S(u|y)\big]\, du \\ &+ \frac{1}{2} m(m-1) \int_{\varphi(y)}^{\infty} \big[\widehat{S}(u|y) - S(u|y)\big]^2 b_y^{m-2}(u)\, du,\end{aligned}$$

where, $\widehat{S}(u|y) \wedge S(u|y) \leq b_y(u) \leq \widehat{S}(u|y) \vee S(u|y)$. So, we obtain:

$$R^y_{m,n} \overset{a.s.}{=} \frac{1}{2} m(m-1) \int_{\varphi(y)}^{\infty} \big[\widehat{S}(u|y) - S(u|y)\big]^2 b_y^{m-2}(u)\, du.$$

By the Law of Iterated Logarithms, we know that $\sup_u \big|\widehat{S}(u|y) - S(u|y)\big| \overset{a.s.}{\leq} C\Big(\frac{\log\log n}{n}\Big)^{1/2}$ for some constant C, so we have

$$\frac{\sqrt{n}}{\sigma(m, y)}\, \big|R^y_{m,n}\big| \overset{a.s.}{\leq} \frac{1}{2}\, \frac{m(m-1)}{\sigma(m, y)}\, \frac{C^2 \log\log n}{\sqrt{n}}\, \int_{\varphi(y)}^{\infty} b_y^{m-2}(u)\, du. \tag{3.8}$$

Let us now analyze the behavior of $\int_{\varphi(y)}^{\infty} b_y^m(u)\, du$ when $m \to \infty$. We can write

$$\int_{\varphi(y)}^{\infty} b_y^m(u)\, du = \int_{\varphi(y)}^{\infty} \big(S(u|y) + r_y(u)\big)^m\, du,$$

for some $r_y(u)$ such that $\widehat{S}(u|y) \wedge S(u|y) - S(u|y) \leq r_y(u) \leq \widehat{S}(u|y) \vee S(u|y) - S(u|y)$. Note that $|r_y(u)| \leq C\Big(\frac{\log\log n}{n}\Big)^{1/2}$. Since $\frac{(S(u|y)+r_y(u))^m - S^m(u|y)}{r_y(u)} = m\left(S(u|y) + g_y(u)\right)^{m-1}$, for some $g_y(u)$ such that $|g_y(u)| \leq |r_y(u)|$, we obtain

$$\begin{aligned}\int_{\varphi(y)}^{\infty} \left(S(u|y) + r_y(u)\right)^m du &\leq \int_{\varphi(y)}^{\infty} S^m(u|y)\, du \\ &+ mC\left(\frac{\log\log n}{n}\right)^{1/2} \int_{\varphi(y)}^{\infty} \left(S(u|y) + g_y(u)\right)^{m-1} du.\end{aligned}$$

Applying the same argument for the exponent $m-1$, one can find

$$\begin{aligned}\int_{\varphi(y)}^{\infty} \left(S(u|y) + g_y(u)\right)^{m-1} du &\leq \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\, du \\ &+ (m-1)C\left(\frac{\log\log n}{n}\right)^{1/2} \int_{\varphi(y)}^{\infty} \left(S(u|y) + h_y(u)\right)^{m-2} du.\end{aligned}$$

for some $h_y(u)$ such that $|h_y(u)| \leq |g_y(u)| \leq |r_y(u)|$. It is clear that

$$\begin{aligned}&\int_{\varphi(y)}^{\infty} \left(S(u|y) + h_y(u)\right)^{m-2} du \leq \int_{\varphi(y)}^{\infty} \left(\widehat{S}(u|y) \vee S(u|y)\right)^{m-2} du \\ &\leq \int_{\varphi(y)}^{\infty} \left(\widehat{S}^{m-2}(u|y) \vee S^{m-2}(u|y)\right) du \leq \int_{\varphi(y)}^{\infty} \widehat{S}^{m-2}(u|y)\, du + \int_{\varphi(y)}^{\infty} S^{m-2}(u|y)\, du.\end{aligned}$$

So, when $m \to \infty$, $\int_{\varphi(y)}^{\infty} \left(S(u|y) + h_y(u)\right)^{m-2} du \overset{a.s.}{=} o(1)$. So finally we obtain when $m \to \infty$,

$$\begin{aligned}\int_{\varphi(y)}^{\infty} b_y^m(u)\, du &\overset{a.s.}{\leq} \left(\varphi_m(y) - \varphi(y)\right) + mC\left(\frac{\log\log n}{n}\right)^{1/2}\left(\varphi_{m-1}(y) - \varphi(y)\right) \\ &\quad + m(m-1)C^2\left(\frac{\log\log n}{n}\right) o(1). \qquad (3.9)\end{aligned}$$

Plugging in (3.8) the results (3.9) and (3.5) and using Lemma 3.1, we obtain for $m \to \infty$,

$$\begin{aligned}\frac{\sqrt{n}}{\sigma(m,y)}\left|R_{m,n}^y\right| &\overset{a.s.}{\leq} \frac{C^2 m^2}{2\sqrt{k_{1,y}} m^{1/2-1/\rho_y}} \frac{\log\log n}{\sqrt{n}} \Bigg\{ \left[\Gamma(1+1/\rho_y)\ell^{-1/\rho_y} + o(1)\right] \\ &\times \left(m^{-1/\rho_y} + mC\big(\frac{\log\log n}{n}\big)^{1/2} m^{-1/\rho_y}\right) + m^2 C^2 \frac{\log\log n}{n}\, o(1)\Bigg\},\end{aligned}$$

so that

$$\begin{aligned}\frac{\sqrt{n}}{\sigma(m,y)}\left|R_{m,n}^y\right| &\overset{a.s.}{\leq} m^{3/2}\frac{\log\log n}{\sqrt{n}}(K_1 + o(1)) + m^{5/2}\frac{(\log\log n)^{3/2}}{n}(CK_1 + o(1)) \\ &\quad + m^{7/2+1/\rho_y}\frac{(\log\log n)^2}{n^{3/2}}\, o(1), \qquad (3.10)\end{aligned}$$

where K_1 is some positive constant. Since under the condition of the theorem $m = m_n = c\, n^{1/3-\varepsilon}\,(\log\log n)^{-2/3}$ all the terms in the r.h.s. of the last inequality converges to 0 when $n \to \infty$, we obtain

$$\frac{\sqrt{n}}{\sigma(m,y)}\, R^y_{m,n} \xrightarrow{a.s.} 0 \ \text{ as } n \to \infty. \tag{3.11}$$

(ii) We now will prove the leading term of (3.7) *converges to a standard normal.* We can rewrite this leading term as

$$\frac{\sqrt{n}\, m}{\sigma(m,y)} \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\big[\widehat{S}(u|y) - S(u|y)\big]\, du = \frac{S_Y(y)}{\widehat{S}_Y(u)} \sum_{i=1}^{n} \frac{W_{n,i}}{\sqrt{n}\,\sigma(m,y)},$$

where $W_{n,i} = \big(m/S_Y(y)\big) \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)\big[\mathbb{1}(X_i \geq u, Y_i \geq y) - S(u|y)\,\mathbb{1}(Y_i \geq y)\big]\, du$. It is easy to see that $\mathbb{E}(W_{n,i}) = 0$ and $\mathbb{V}(W_{n,i}) = \sigma^2(m,y)$. By the Lindberg-Feller theorem (Serfling, 1980, p. 29) we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{W_{n,i}}{\sigma(m,y)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \ \text{ as } n \to \infty, \tag{3.12}$$

under the Liapounoff condition, *i.e.* if

$$\frac{n\mathbb{E}\big(|W_{n,i}|^3\big)}{\big[n\mathbb{V}(W_{n,i})\big]^{3/2}} \longrightarrow 0, \ \text{ as } n \to \infty. \tag{3.13}$$

The Liapounoff condition is easy to check under the assumptions of the theorem. Indeed, $\mathbb{E}\big(|W_{n,i}|^3\big) = \mathbb{E}\big(W^2_{n,i}|W_{n,i}|\big)$ and since $\big|\mathbb{1}(X_i \geq u, Y_i \geq y) - S(u|y)\,\mathbb{1}(Y_i \geq y)\big| \leq 1$, we have

$$|W_{n,i}| \leq \frac{m}{S_Y(y)} \int_{\varphi(y)}^{\infty} S^{m-1}(u|y)du = \frac{m}{S_Y(y)}\big(\varphi_{m-1}(y) - \varphi(y)\big).$$

So, $\mathbb{E}\big(|W_{n,i}|^3\big) \leq \big(m/S_Y(y)\big)\big(\varphi_{m-1}(y) - \varphi(y)\big)\sigma^2(m,y)$ and we obtain

$$\frac{n\mathbb{E}\big(|W_{n,i}|^3\big)}{\big[n\mathbb{V}(W_{n,i})\big]^{3/2}} \leq \frac{m}{\sqrt{n}\, S_Y(y)}\, \frac{\varphi_{m-1}(y) - \varphi(y)}{\sigma(m,y)}.$$

Under the regularity condition (2.5), Lemma 3.1 and (3.5), we have, as $m \to \infty$, $\sigma^2(m,y) \geq k_{1,y} m^{1-2/\rho_y}$ and $\varphi_{m-1}(y) - \varphi(y) \sim \Gamma(1 + 1/\rho_y)\big(\frac{1}{\ell_y(m-1)}\big)^{1/\rho_y}$, so that

$$\frac{n\mathbb{E}\big(|W_{n,i}|^3\big)}{\big[n\mathbb{V}(W_{n,i})\big]^{3/2}} \leq K_2 \frac{m^{1/2}}{\sqrt{n}},$$

where K_2 is some poistive constant. The r.h.s. of the latter inequality tends to zero if $n \to \infty$ and $m \to \infty$ such that $m/n \to 0$ which is the case for the sequence $m = m_n$ given in the assumption of the theorem. Finally, since $\big(S_Y(y)/\widehat{S}_Y(y)\big) \xrightarrow{a.s.} 1$, as $n \to \infty$, we obtain the desired result. $\square$

Rate of convergence

It is interesting to analyze the resulting rate of convergence of the estimator as a function of n. We have as $n \to \infty$, $\tau_n\big(\hat{\varphi}_m(y) - \varphi_m(y)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ with $\tau_n = \sqrt{n}/\sigma(m,y)$ and $m = m_n = c\, n^{1/3-\varepsilon}\,(\log\log n)^{-2/3}$. We know by Lemme 3.1 that as $n \to \infty$,

$$k_{1,y} c^{1-2/\rho_y} n^{(1/3-\varepsilon)(1-2/\rho_y)-1} (\log\log n)^{-(2/3)(1-2/\rho_y)} \le \tau_n^{-2} \le k_{2,y} c^{2-2/\rho_y} n^{(1/3-\varepsilon)(2-2/\rho_y)-1} (\log\log n)^{-(2/3)(2-2/\rho_y)}.$$

We remember that $\rho_y = \beta_y + q + 1$, where $q \ge 1$ and $\beta_y > -1$ (see the discussion after (2.7) above). In the particular case where the extreme value index $\rho_y \ge 2$ we get as $n \to \infty$

$$c_1 n^{-(1/3)(1-1/\rho_y)+1/2} (\log\log n)^{(1/3)(2-2/\rho_y)} \le \tau_n \le c_2 n^{1/2} (\log\log n)^{(1/3)(1-2/\rho_y)}.$$

This case is of particular interest when the joint density of (X,Y) has a jump at the frontier (i.e. $\beta_y = 0$, an often used assumption in the econometric literature). We have clearly in this case as $q \downarrow 1$,

$$c_1\big(n \log\log n\big)^{1/3} \le \tau_n \le c_2 n^{1/2},$$

and as $q \uparrow \infty$,

$$c_1 n^{1/6} (\log\log n)^{2/3} \le \tau_n \le c_2 n^{1/2} (\log\log n)^{1/3}.$$

So, even if the data dimension explodes, the convergence rate does not deteriorate too much avoiding thus, in a sense and partly, the "curse of dimensionality" that is typical of many nonparametric estimators. This comes from Cazals et al. (2002) where it is shown that the order-m frontier is a linear functional of a survival function. Since the survival function $S(y|X \le x)$ is estimated by its empirical version at the $\sqrt{n}$ rate for any dimension of x, the order-m frontier estimator keeps this rate, for fixed m. We loose something of this, but not all, when $m = m_n$ is increasing slowly enough to infinity, as $n \to \infty$.

3.2 Estimation of the frontier $\varphi(y)$

Since $\varphi_m(y) \to \varphi(y)$ as $m \to \infty$, the result of the preceding section can be used to define an estimator of the "full" frontier itself. From Theorem 3.1, if $m_n < n^{1/3}(\log\log n)^{-2/3}$, we have

$$\frac{\sqrt{n}}{\sigma(m_n,y)}\big(\hat{\varphi}_{m_n}(y) - \varphi(y) - B_{m_n}(y)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \tag{3.14}$$

where from (3.5),

$$B_{m_n}(y) = \varphi_{m_n}(y) - \varphi(y) = \Gamma\left(1 + \frac{1}{\rho_y}\right)\left(\frac{1}{m_n\,\ell_y}\right)^{1/\rho_y} + o(m_n^{-1/\rho_y}). \tag{3.15}$$

We see that the value of the bias introduced by using the partial order-m_n frontier to estimate the full frontier is bounded below $\big(\sqrt{n}/\sigma(m_n,y)\big)B_{m_n}(y) > K_3 n^{1/3}(\log\log n)^{1/3}$ for some constant K_3, and this does not vanish when $n \to \infty$.

So, in practice for large values of n (and so of m), we will rather use the following asymptotic approximation:

$$\hat{\varphi}_m(y) - \varphi(y) \approx \mathcal{N}(B_m(y), \frac{\sigma^2(m,y)}{n}), \tag{3.16}$$

where for doing practical inference $B_m(y)$ and $\sigma(m,y)$ have to be consistently estimated. A consitent estimator of $\sigma(m,y)$ is provided by a plugging version of (3.2), whereas, a consistent estimator of $B_m(y)$ can be obtained through the leading part of (3.15) once ρ_y and ℓ_y are known or consistently estimated. The next section suggests a way for estimating these two parameters, using the properties of order-m frontiers.[4]

4 Consistent estimators of the Bias

4.1 Consistent estimators of ρ_y and ℓ_y

We will use here an approach inspired by the classical Pickands tail index estimator, analyzed and developed in our frontier setup in Daouia et al. (2010). The Pickands estimator is based on comparing different quantile-type estimators of the frontier. As well known from the literature, and illustrated in Daouia et al., the estimator is rather unstable and provide disappointing results unless the sample size is larger than, say 1000. Daouia et al. (2010) also analyze a moment type of estimator providing slightly better behavior in moderated sample sizes (say larger than 500).

In this paper, we adapt the approach by using the order-m estimator of the frontier instead of the order-α quantile estimator of the frontier. Indeed, when considering the asymptotic expression for $\varphi_m(y) - \varphi(y)$ given by (3.5) for the values m, am and a^2m, where a is some fixed integer with $a \geq 2$, we see that

$$\lim_{m\to\infty} \frac{\varphi_m(y) - \varphi_{am}(y)}{\varphi_{am}(y) - \varphi_{a^2m}(y)} = a^{1/\rho_y}.$$

This suggests the following estimator

$$\hat{\rho}_y = \log(a) \left\{ \log\Big(\frac{\hat{\varphi}_{m_n}(y) - \hat{\varphi}_{am_n}(y)}{\hat{\varphi}_{am_n}(y) - \hat{\varphi}_{a^2m_n}(y)} \Big) \right\}^{-1}. \tag{4.1}$$

[4]It is well known that the order-m (and order-α) frontier estimates could be non-monotone. Monotonic versions can be easily obtained by isotonizing the unconstrained estimates. Daouia and Simar (2005) investigate this approach and show that all the nice properties of the original estimators are maintained after isotonization.

It is also easy to see that

$$\lim_{m\to\infty}\frac{1}{m}\left[\frac{\Gamma(1+1/\rho_y)\big(1-a^{-1/\rho_y}\big)}{\varphi_m(y)-\varphi_{am}(y)}\right]^{\rho_y}=\ell_y,$$

that can lead to the estimator of ℓ_y

$$\hat{\ell}_y=\frac{1}{m_n}\left[\frac{\Gamma(1+1/\hat{\rho}_y)\big(1-a^{-1/\hat{\rho}_y}\big)}{\hat{\varphi}_{m_n}(y)-\hat{\varphi}_{am_n}(y)}\right]^{\hat{\rho}_y}. \tag{4.2}$$

The consistency of these estimators is provided by the following theorems.

Theorem 4.1. *Under the regularity conditions of Theorem 3.1,*

$$\hat{\rho}_y\xrightarrow{P}\rho_y \text{ and } \hat{\ell}_y\xrightarrow{P}\ell_y \text{ as } n\to\infty, \tag{4.3}$$

for any y such that $S_Y(y)>0$,

Proof: By Theorem 3.1, we have $\hat{\varphi}_m(y)-\varphi_m(y)=O_p\big(\sigma(m,y)/\sqrt{n}\big)$. Now, by (3.5), and by Lemma 3.1, we obtain

$$\hat{\varphi}_m(y)-\varphi(y)=C_y\left(\frac{1}{m}\right)^{1/\rho_y}+o\big(m^{-1/\rho_y}\big)+O_p\Big(\frac{m^{1-1/\rho_y}}{\sqrt{n}}\Big)$$

where $C_y=\Gamma\left(1+\frac{1}{\rho_y}\right)\left(\frac{1}{\ell_y}\right)^{1/\rho_y}$. Similarly we have for all $a\geq 2$

$$\begin{aligned}\hat{\varphi}_{am}(y)-\varphi(y)&=C_y\left(\frac{1}{am}\right)^{1/\rho_y}+o\big(m^{-1/\rho_y}\big)+O_p\Big(\frac{m^{1-1/\rho_y}}{\sqrt{n}}\Big)\\ \hat{\varphi}_{a^2m}(y)-\varphi(y)&=C_y\left(\frac{1}{a^2m}\right)^{1/\rho_y}+o\big(m^{-1/\rho_y}\big)+O_p\Big(\frac{m^{1-1/\rho_y}}{\sqrt{n}}\Big).\end{aligned}$$

Now by doing the differences we have

$$\begin{aligned}m^{1/\rho_y}\big(\hat{\varphi}_m(y)-\hat{\varphi}_{am}(y)\big)&=C_y\big(1-1/a^{1/\rho_y}\big)+o(1)+O_p\Big(\frac{m}{\sqrt{n}}\Big)\\ (am)^{1/\rho_y}\big(\hat{\varphi}_{am}(y)-\hat{\varphi}_{a^2m}(y)\big)&=C_y\big(1-1/a^{1/\rho_y}\big)+o(1)+O_p\Big(\frac{m}{\sqrt{n}}\Big),\end{aligned}$$

leading to

$$\frac{\hat{\varphi}_m(y)-\hat{\varphi}_{am}(y)}{\hat{\varphi}_{am}(y)-\hat{\varphi}_{a^2m}(y)}=a^{1/\rho_y}\frac{C_y\big(1-1/a^{1/\rho_y}\big)+o(1)+O_p\Big(\frac{m}{\sqrt{n}}\Big)}{C_y\big(1-1/a^{1/\rho_y}\big)+o(1)+O_p\Big(\frac{m}{\sqrt{n}}\Big)}.$$

As $m/\sqrt{n} \to 0$ as $n \to \infty$, the ratio on the right hand side converges in probability to 1, so that

$$\frac{\hat{\varphi}_m(y) - \hat{\varphi}_{am}(y)}{\hat{\varphi}_{am}(y) - \hat{\varphi}_{a^2m}(y)} \xrightarrow{P} a^{1/\rho_y},$$

which gives $\hat{\rho}_y \xrightarrow{P} \rho_y$. On the other hand, since

$$m^{1/\rho_y}\big(\hat{\varphi}_m(y) - \hat{\varphi}_{am}(y)\big) = \Gamma(1+1/\rho_y)(1-1/a^{1/\rho_y})(\ell_y)^{-1/\rho_y} + o(1) + O_p\Big(\frac{m}{\sqrt{n}}\Big),$$

we have by using $m/\sqrt{n} \to 0$ and $\hat{\rho}_y \xrightarrow{P} \rho_y$ as $n \to \infty$,

$$\frac{1}{m}\left[\frac{\Gamma(1+1/\hat{\rho}_y)\big(1-a^{-1/\hat{\rho}_y}\big)}{\hat{\varphi}_m(y) - \hat{\varphi}_{am}(y)}\right]^{\hat{\rho}_y} \xrightarrow{P} \ell_y,$$

which gives $\hat{\ell}_y \xrightarrow{P} \ell_y$. $\square$

4.2 Practical choice of a and m

The choice of optimal values of a and m is an open theoretical issue. Nothing guarantees that an optimal choice of m for estimating the frontier function $\varphi(y)$ is the same as the selected optimal m for estimating the regularity parameters ρ_y and ℓ_y. Moreover, the selection optimal values for m and a depend heavily on the tail index ρ_y. Daouia et al. (2010) faced similar issues for choosing the extreme order $\alpha = \alpha_n$ for providing the quantile-based regularized version of the FDH estimator. This topic is in our research agenda, but we know from the literature on extreme-value theory that the determination of an optimal sample fraction for the estimation of the tail index ρ_y is not an easy task.

However, in practice, we will see from the simulations below that the performances of our estimators are much less sensitive to the choice of a and m than in the quantile-based approach of Daouia et al. (2010). This is probably due to the construction of the order-m_n frontier estimator as a linear combination of large order statistics, while the order-α_n frontier estimator is given by a single extreme order statistic (see, *e.g.*, Daouia and Gijbels (2011a) for explicit formulations of both empirical partial frontiers in terms of conditional order statistics). In most of our illustrations, we have chosen $m_n = N_y^{1/3}$, where $N_y = \sum_{i=1}^n \mathbb{1}(Y_i \geq y)$ is the number of observations with $Y_i \geq y$. This choice guarantees by Theorem 3.1 the regular behavior of the estimator $\hat{\varphi}_{m_n}(y)$ as $n \to \infty$ and as seen above, it guarantees also the consistency of the estimators $\hat{\rho}_y$ and $\hat{\ell}_y$.

The selection of $a \geq 2$ is much less important: the results are rather stable relative to this choice. Higher values of a will give more weights to extreme data points and this choice

typically depends on the value of the tail index itself. However, in all the Monte-Carlo experiments below, it turns out that the choice $a = 10$ provides quite reasonable estimates with a nice behavior of the estimators in terms of Bias and Mean Squared Errors (MSE) for both the regularity parameters as well as for the frontier itself. In some cases, where the tail index is higher, smaller values may be preferable. We will give in Subsections 5.1 and 5.2 the Monte-Carlo results for the Bias and MSE with the quite different values $a = 2$ and $a = 10$. We will see that the resulting Bias and MSE are much less sensitive to the choice of a than to the choice of the order α in the extreme quantile-based approach, and globally, as commented below, our results are much better by a significative order of magnitude.

When working with particular samples of real data, and for the estimation of ρ_y, we have to tune the choice of a and m more carefully to obtain sensible results and to avoid numerical problems in solving (4.1) (see the real data example in Paragraph 5.2.3).

For the final evaluation of the confidence intervals for $\varphi(y)$, we use the normal approximation centered at $\widetilde{\varphi}_m(y)$, the bias-corrected order-m estimate defined as

$$\widetilde{\varphi}_m(y) = \widehat{\varphi}_m(y) - \widehat{B}_m(y), \tag{4.4}$$

where $\widehat{B}_m(y)$ is the plug-in version of $B_m(y)$, replacing ρ_y and ℓ_y by their consistent estimators derived above. We know from (3.16) that an asymptotic $(1-\alpha)\times 100\%$ confidence interval for $\varphi(y)$ is given by

$$\varphi(y) \in \left[\widehat{\varphi}_m - B_m(y) \pm z_{1-\alpha/2}\frac{\sigma(m,y)}{\sqrt{n}}\right], \tag{4.5}$$

where $z_{1-\alpha/2}$ stands for the $(1-\alpha/2)$th quantile of the standard normal distribution. Then we estimate this interval by plugging the consistent estimator of $B_m(y)$. Since by doing so, we will increase the variance of the estimator $\hat{\varphi}_m(y)$, we adjust the variance $\sigma(m,y)/\sqrt{n}$ by a bootstrap estimator of the standard deviation of the bias-corrected estimator $\widetilde{\varphi}_m(y)$. We will see below, via a Monte-Carlo experiment, that the achieved coverages of these estimates of confidence intervals are quite reasonable.

4.3 The unregularized order-m estimator

An alternative to our regularized robust estimator of the frontier is the unregularized order-m estimator proposed in Cazals et al. (2002), i.e. the order-m with $m = m_n \to \infty$ fast enough when $n \to \infty$. By making use of (2.14) we could employ the quantiles of the Weibull distribution to build alterantive estimates of confidence intervals for $\varphi(y)$. Note that here too, we have to estimate ρ_y and ℓ_y for getting these intervals.

From Theorem 3.2 in Cazals et al. (2002) we know that m_n should be of the order $O(CnS_Y(y)\log(n))$, where $C > 1/(q+1)$ to obtain the Weibull approximation for

$\big(\hat{\varphi}_{m_n}(y) - \varphi(y)\big)$, but this is not very helpful to choose m_n in practice. We know also that if m_n is too large it will coincide with the FDH estimator, loosing all its robustness properties. To the best of our knowledge, only Daouia and Gijbels (2011b) have proposed a 'semi-automatic' procedure for selecting appropriate values for m_n, but there is no complete automatic data driven technique that would be useful in a Monte-Carlo setup. When choosing in our simulations $m_n = N_y \log(N_y)/(q+1)$, the order-$m$ frontier estimator was confounded with the FDH estimator in most of the cases, with very poor robustness properties. So we selected $m_n = N_y$ in the examples below (to be contrasted with the choice $m_n = N_y^{1/3}$ for our regularized estimator).

4.4 Asymptotic normality of $\tau_n\big(\widetilde{\varphi}_m(y) - \varphi(y)\big)$

We have proved under some regularity conditions and with a suitable choice of the sequence $m = m_n$, that $\tau_n\big(\widehat{\varphi}_m(y) - \varphi(y)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$, with $\tau_n = \sqrt{n}/\sigma(m,y)$. The question is now to analyze the asymptotic behavior of the bias-corrected order-m estimator defined in (4.4), i.e., the behavior of $\tau_n\big(\widetilde{\varphi}_m(y) - \varphi(y)\big)$ and to characterize the extra conditions under which this expression also converges to a standard normal distribution. We have imposed that the survivor function S satisfies the first order regularity condition (2.5) which determines the property (3.5). We shall need a second order refinement of this relation which reads as follows: there exists $\alpha_y > 0$ such that

$$\varphi_m(y) - \varphi(y) = \Gamma\left(1 + \frac{1}{\rho_y}\right)\left(\frac{1}{m\,\ell_y}\right)^{1/\rho_y} + o(m^{-(1+\alpha_y)/\rho_y}). \tag{4.6}$$

Here we impose the reminder term to be $o(m^{-(1+\alpha_y)/\rho_y})$ which is supposed to be $o(m^{-1/\rho_y})$ in the first order regularity condition (i.e. $\alpha_y = 0$). So, (3.5) motivates the stringent condition (4.6), which is used in the next theorem to get the asymptotic normality, assuming that ρ_y is given with ℓ_y being estimated by using the value $m = \widetilde{m}$ described in the theorem.

Theorem 4.2. *Under the conditions* (2.5) *and* (4.6) *with* $\alpha_y > \rho_y$, *if* $m = cn^{1/3-\varepsilon}$ *with* $0 < \varepsilon < \frac{1}{3}$, *and* $\widetilde{m} = \tilde{c}n^{1/3-\tilde{\varepsilon}}$ *for some* $\tilde{\varepsilon} > \frac{1}{2}\varepsilon + \frac{1}{6}$ *such that* $\frac{\alpha_y}{\rho_y}\big(\frac{1}{3} - \tilde{\varepsilon}\big) + \frac{1}{2}\big(\frac{1}{3} - \varepsilon\big) - \frac{1}{2} > 0$, *we have for any* y *such that* $S_Y(y) > 0$,

$$\tau_n\big(\widetilde{\varphi}_m(y) - \varphi(y)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \quad as \quad n \to \infty, \tag{4.7}$$

where $\widetilde{\varphi}_m(y) = \hat{\varphi}_m(y) - \widehat{B}_m(y)$ *and* $\widehat{B}_m(y) = \Gamma\left(1 + \frac{1}{\rho_y}\right)\left(\frac{1}{m\,\hat{\ell}_{\widetilde{m},y}}\right)^{1/\rho_y}$.

Proof: The result is true if

$$\frac{\sqrt{n}}{\sigma(m,y)}\big(\widehat{B}_m(y) - (\varphi_m(y) - \varphi(y))\big) \xrightarrow{P} 0,$$

or equivalently if

$$\frac{\sqrt{n}}{\sigma(m,y)}\left(\widehat{B}_m(y)-\Gamma\left(1+\frac{1}{\rho_y}\right)\left(\frac{1}{m\,\ell_y}\right)^{1/\rho_y}+o\big(m^{-(1+\alpha_y)/\rho_y}\big)\right)\xrightarrow{P}0. \tag{4.8}$$

(i) Let us consider the last term of (4.8) where $m\sim n^{-1/3}$ (it is not hard to verify that we can neglect the log log term and the ε appearing in the definition of $m=m_n$ in Theorem 3.1, without loss of generality). Using the bounds given in Lemma 3.1, it is clear that this term is $o(1)$ if $\alpha_y>\rho_y$.

(ii) For the first term of (4.8), we have

$$\begin{aligned}&\frac{\sqrt{n}}{\sigma(m,y)}\left(\widehat{B}_m(y)-\Gamma\left(1+\frac{1}{\rho_y}\right)\left(\frac{1}{m\,\ell_y}\right)^{1/\rho_y}\right)\\&\qquad=\frac{\sqrt{n}}{\sigma(m,y)}\Gamma\Big(1+\frac{1}{\rho_y}\Big)\frac{1}{m^{1/\rho_y}}\left[\Big(\frac{1}{\hat{\ell}_{\widetilde{m},y}}\Big)^{1/\rho_y}-\Big(\frac{1}{\ell_y}\Big)^{1/\rho_y}\right],\end{aligned}$$

where $\hat{\ell}_{\widetilde{m},y}=\frac{1}{\widetilde{m}}\left[\frac{\Gamma\left(1+1/\rho_y\right)\left(1-a^{-1/\rho_y}\right)}{\hat{\varphi}_{\widetilde{m}}(y)-\hat{\varphi}_{a\widetilde{m}}(y)}\right]^{\rho_y}$. Let us define similarly $\ell_{\widetilde{m},y}=\frac{1}{\widetilde{m}}\left[\frac{\Gamma\left(1+1/\rho_y\right)\left(1-a^{-1/\rho_y}\right)}{\varphi_{\widetilde{m}}(y)-\varphi_{a\widetilde{m}}(y)}\right]^{\rho_y}$. Then we can write the first term of (4.8) as

$$\begin{aligned}&\frac{\sqrt{n}}{\sigma(m,y)}\Gamma\Big(1+\frac{1}{\rho_y}\Big)\frac{1}{m^{1/\rho_y}}\left[\Big(\frac{1}{\hat{\ell}_{\widetilde{m},y}}\Big)^{1/\rho_y}-\Big(\frac{1}{\ell_{\widetilde{m}y}}\Big)^{1/\rho_y}\right]\\&\qquad+\frac{\sqrt{n}}{\sigma(m,y)}\Gamma\Big(1+\frac{1}{\rho_y}\Big)\frac{1}{m^{1/\rho_y}}\left[\Big(\frac{1}{\ell_{\widetilde{m},y}}\Big)^{1/\rho_y}-\Big(\frac{1}{\ell_y}\Big)^{1/\rho_y}\right]\\&=I+II.\end{aligned} \tag{4.9}$$

Now the first term I can be written as

$$\begin{aligned}I&=\frac{\sqrt{n}}{\sigma(m,y)}\frac{1}{m^{1/\rho_y}}\frac{\widetilde{m}^{1/\rho_y}}{(1-a^{-1/\rho_y})}\left[\big(\hat{\varphi}_{\widetilde{m}}(y)-\varphi_{\widetilde{m}}(y)\big)-\big(\hat{\varphi}_{a\widetilde{m}}(y)-\varphi_{a\widetilde{m}}(y)\big)\right]\\&=\frac{\sqrt{n}}{\sigma(m,y)}\frac{\widetilde{m}^{1/\rho_y}}{m^{1/\rho_y}}\frac{\sigma(\widetilde{m},y)}{\sqrt{n}}O_P(1)=O_P\left(\frac{\sigma(\widetilde{m},y)}{\sigma(m,y)}\,\frac{\widetilde{m}^{1/\rho_y}}{m^{1/\rho_y}}\right).\end{aligned}$$

By Lemma 3.1 we have

$$\frac{\sigma(\widetilde{m},y)}{\sigma(m,y)}\,\frac{\widetilde{m}^{1/\rho_y}}{m^{1/\rho_y}}\le\left(\frac{k_{2,y}}{k_{1,y}}\right)^{1/2}\frac{\widetilde{m}}{m^{1/2}},$$

where the last ratio tends to zero as soon as $2\tilde{\varepsilon}>\varepsilon+1/3$. Therefore we have $I=o_P(1)$.

Let us now turn to the second term II of (4.9). We see that

$$II=\frac{\sqrt{n}}{\sigma(m,y)}\Gamma\Big(1+\frac{1}{\rho_y}\Big)\frac{1}{m^{1/\rho_y}}\left[\frac{\widetilde{m}^{1/\rho_y}(\varphi_{\widetilde{m}}(y)-\varphi_{a\widetilde{m}}(y))}{\Gamma\big(1+1/\rho_y\big)\big(1-a^{-1/\rho_y}\big)}-\Big(\frac{1}{\ell_y}\Big)^{1/\rho_y}\right],$$

which by using our regularity condition (4.6) can be written as

$$II = \frac{\sqrt{n}}{\sigma(m,y)}\Gamma\Big(1+\frac{1}{\rho_y}\Big)\frac{1}{m^{1/\rho_y}}\left[\widetilde{m}^{1/\rho_y}o\big(\widetilde{m}^{-(1+\alpha_y)/\rho_y}\big)\right] = \frac{\sqrt{n}}{\sigma(m,y)}\ o\Big(\frac{\widetilde{m}^{-\alpha_y/\rho_y}}{m^{1/\rho_y}}\Big).$$

By using again Lemma 3.1, we have

$$|II| < \frac{\sqrt{n}}{m^{1/2-1/\rho_y}}\frac{\widetilde{m}^{-\alpha_y/\rho_y}}{m^{1/\rho_y}}o(1).$$

So, II will be $o(1)$ if $\dfrac{\sqrt{n}}{m^{1/2}}\dfrac{1}{\widetilde{m}^{\alpha_y/\rho_y}}$ is $O(1)$. It is easily seen that this happens with the selected sequences m and $\widetilde{m}$ as soon as

$$\frac{\alpha_y}{\rho_y}\big(\frac{1}{3}-\tilde{\varepsilon}\big)+\frac{1}{2}\big(\frac{1}{3}-\varepsilon\big)-\frac{1}{2}>0.$$

This completes the proof of the theorem. □

Remark 4.1. *For the conditions on the regularization parameters appearing in the theorem to be satisfied, it is enough to have:*

$$0<\varepsilon<\frac{1}{3},\quad \alpha_y>2\rho_y,$$
$$\frac{1}{6}+\frac{1}{2}\varepsilon<\tilde{\varepsilon}<\frac{1}{3}-\frac{1}{2(1+\alpha_y/\rho_y)}.$$

The proof also shows that the difficulty does not come from the estimation of ℓ_y (driven by part I of (4.9) where it is required to select $\widetilde{m} = o(m^{1/2})$), but only from the bias approximation in (3.5), which has an error of $o\big(m^{-(1+\alpha_y)/\rho_y}\big)$ (driven by the last term of (4.8) and part II of (4.9)).

The last point to consider is the relevance of the regularity condition (4.6). This property is a condition on $S(x|y)$ we may detail. First let us rewrite $S(x|y)$ as $e^{-\Lambda_y(x)}$, with $\Lambda_y(\cdot)$ being the conditional integrated hazard function associated with $S(\cdot|y)$. Then starting from (2.10) and using elementary algebra, we obtain

$$\begin{aligned}\varphi_m(y)-\varphi(y) &= \int_{\varphi(y)}^{\infty} S^m(x|y)\,dx = \int_{\varphi(y)}^{\infty} e^{-m\Lambda_y(x)}\,dx \\ &= \int_0^{\infty} e^{-u}\Lambda_y^{-1'}\Big(\frac{u}{m}\Big)\frac{1}{m}\,du = \int_0^{\infty} e^{-u}\Lambda_y^{-1}\Big(\frac{u}{m}\Big)\,du. \qquad (4.10)\end{aligned}$$

The last equality is obtained under the condition $\lim\limits_{u\to\infty} e^{-u}\Lambda_y^{-1}(\dfrac{u}{m}) = 0$, where Λ_y^{-1} stands for the inverse function of Λ_y.

If for instance, $S(x|y)$ is a Weibull distribution, $\Lambda_y^{-1}(t) = \left(\frac{1}{\ell_y}\right)^{1/\rho_y} t^{1/\rho_y}$. This shows that in the Weibull case, α_y is infinite: our approximation of bias is exact. A more general assumption will be

$$\Lambda_y^{-1}(t) = \left(\frac{1}{\ell_y}\right)^{1/\rho_y} t^{1/\rho_y} + c_y t^{(1+\alpha_y)/\rho_y} + o\big(t^{(1+\alpha_y)/\rho_y}\big) \quad \text{as} \quad t \downarrow 0, \tag{4.11}$$

for some constant $c_y \in \mathbb{R}$. Under the property (4.11), it can be easily seen by using (2.10) that the second order condition (4.6) holds with $c_y = 0$. The first order condition (2.5) says that this property is satisfied for $\alpha_y = 0$. The asymptotic normality of $\tau_n\big(\widetilde{\varphi}_m(y) - \varphi(y)\big)$ obtained in Theorem 4.2 requires a stronger condition relating α_y, ρ_y, m and $\widetilde{m}$.

It is also interesting to remark that (4.6) or the more general condition (4.11) parallels in fact the well-known extreme-value condition required to prove the asymptotic normality of the bias-corrected frontier estimator in the quantile-based framework. Indeed, as pointed out in footnote 3, under the sufficient first order condition (2.5) and when the order $\alpha = \alpha_n \to 1$ slowly enough as $n \to \infty$, Daouia et al. (2010) established the asymptotic normality of the unregularized estimator $\widehat{\varphi}_\alpha(y) := \inf\{x|\widehat{S}_{X|Y}(x|y) < \alpha\}$ of the partial quantile-type frontier function $\varphi_\alpha(y) := \inf\{x|S_{X|Y}(x|y) < \alpha\}$ (this parallels our Theorem 3.1 for order-m partial frontiers). Then for estimating the full frontier function itself $\varphi(y)$, the underlying idea was to shift the *anchor* order-α partial frontier to the right place. To derive a normal limiting distribution for the resulting regularized estimator of $\varphi(y)$, the following second order condition is required on the quantile-type funcion

$$\varphi_{1-\frac{t}{S_Y(y)}}(y) = c_0 + c_1\rho_y(t^{1/\rho_y} - 1) + c_2 t^{(1+\alpha_y)/\rho_y} + o\left(t^{(1+\alpha_y)/\rho_y}\right) \quad \text{as} \quad t \downarrow 0, \tag{4.12}$$

for some constants $c_0 \in \mathbb{R}$, $c_1 > 0$, $c_2 \neq 0$ and $\alpha_y > 0$. We refer to *e.g.* Ferreira, de Haan and Peng (2003) for a detailed motivation of the traditional extreme-value condition (4.12) in the non-conditional case.

Thus, as may be seen from (4.11) and (4.12), we shall have to require a similar extra condition on the tail-quantile function $\varphi_{\alpha_t}(y)$, with $\alpha_t = e^{-t}$ for the order-m frontier modeling (since $\Lambda_y^{-1}(t)$ is identical to $\varphi_{e^{-t}}(y)$) and $\alpha_t = 1 - \frac{t}{S_Y(y)}$ for the order-α frontier modeling.

Finally, we note that we restrict ourselves in Theorem 4.2 to the particular case where ρ_y is known. This corresponds at least to the usual assumption in the econometric literature on nonparametric frontier analysis that the density of data has jumps at the frontier, or equivalently $\rho_y = q + 1$ (see the discussion below (2.7)). We do not enter here into the case where ρ_y is unknown. The question of whether the asymptotic normality in Theorem 4.2 also holds when replacing ρ_y by its estimator $\hat{\rho}_y$ is a topic of interest for future research.

5 Illustrative Examples

5.1 Some Monte-Carlo experiments

To facilitate the comparison with the results obtained in Daouia et al. (2010), we have chosen in the illustrations the output orientation[5]. Here, the bias-corrected regularized estimator is given by $\widetilde{\varphi}_m(x) = \hat{\varphi}_m(x) + \widehat{B}_m(x)$. In the illustrations below, we limit the presentation to the case of one-input and one output. Multivariate extensions are immediate (multi-inputs for a production function with one output and multi-outputs for a cost or a one-dimensional input function). Since our estimator does not suffer too much from the curse of dimensionality (see the discussion above), we limit the presentation to the bivariate case where nice figures perfectly illustrate the performance of our estimator.

5.1.1 Uniform distribution

We first simulate, as in Daouia et al., random samples (X_i, Y_i), $i = 1, \ldots, n$ uniformly distributed on the triangle limited by the frontier $\varphi(x) = x$ with $0 \leq x \leq 1$. Table 1 displays the results. The estimation is performed for $x = 1$, so that the sample size n coincides with the "effective" sample size N_x, the number of observations at the left of $x = 1$. We computed also the estimators with the known true value of ρ, which is $\rho_0 = 2$ in this example.

We observe a nice behavior of our estimators, with an increasing accuracy, as expected, when the effective sample size N_x increases. The estimation of ρ and ℓ is not an easy task, but still we have a reasonable behavior, with the simple rule we have chosen for m and a: $m = N_x^{1/3}$ and $a = 10$. The estimator $\widetilde{\varphi}_m$ has a very nice behavior for all values of N_x and we note that it has much better properties than the usual FDH estimator $\hat{\varphi}$ (in terms of both bias and mean squared error) and than the unregularized estimator of the frontier from Cazals et al. (2002) (called the CFS estimator hereafter and noted $\tilde{\varphi}_{CFS}$ in the Table). We see that with the smallest possible value for a, $(a = 2)$, the performances are less good but still comparable or even better than the CFS.

The cost of estimating ρ (which in most econometric applications is supposed to be equal to $p + 1$, i.e. there is a jump of the joint density of (X, Y) at the frontier) appears clearly when comparing the results $\widetilde{\varphi}_m(\rho_0)$ for the estimation of the frontier when the true value of $\rho = 2$ is known: they are much better and much less sensitive to the choice of a.

[5] We can find in the appendix the change of the notations.

Table 1: *Bias and Mean Squared Error (MSE) of the estimates over 1000 Monte-Carlo simulations: Uniform case, true values are* $\varphi_0 = 1$, $\rho_0 = 2$ *and* $\ell_0 = 1$

	$N_x = 100$		$N_x = 500$		$N_x = 1000$		$N_x = 5000$	
$a = 2$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	0.628002	1.555049	0.353080	0.274196	0.278767	0.154348	0.147660	0.043355
$\hat{\ell}$	0.399147	0.380958	0.321304	0.135647	0.286953	0.104483	0.202129	0.052757
$\hat{\ell}(\rho_0)$	0.504116	0.362388	0.251523	0.083483	0.188782	0.046584	0.104923	0.014347
$\tilde{\varphi}$	0.041742	0.024328	0.022050	0.002758	0.016362	0.001210	0.006670	0.000218
$\tilde{\varphi}(\rho_0)$	-0.046530	0.004700	-0.020102	0.000910	-0.013962	0.000432	-0.006043	0.000088
$a = 10$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	-0.423441	0.405546	-0.086590	0.167862	-0.030069	0.103272	0.006700	0.032857
$\hat{\ell}$	0.129996	0.143781	0.134017	0.140254	0.123555	0.109258	0.090057	0.052087
$\hat{\ell}(\rho_0)$	0.436484	0.357537	0.183585	0.063740	0.129845	0.031946	0.071624	0.009915
$\tilde{\varphi}$	-0.070644	0.008321	-0.017592	0.001231	-0.008925	0.000501	-0.002371	0.000087
$\tilde{\varphi}(\rho_0)$	-0.035497	0.003719	-0.011318	0.000603	-0.006952	0.000255	-0.002778	0.000053
$\tilde{\varphi}_{CFS}$	-0.113433	0.014708	-0.050186	0.002902	-0.035234	0.001408	-0.016076	0.000294
$\hat{\varphi}$	-0.090498	0.010401	-0.040257	0.002071	-0.028140	0.000993	-0.012811	0.000206

Finally, by looking to Tables 1 and 3 in Daouia et al. (2010) (and the reference [3] given there) using also Pickands estimator of ρ, but with extreme quantile-type frontiers, we see that we obtain here much more accurate estimators of both ρ and φ. To summarize this comparison, we present some limited results from Daouia et al., in Table 2: we have the same scenario, with comparable sample sizes. In this table we provide the MSE with the **optimal** choice of the quantile order α_n^* in each case. In the original full tables in Daouia et al. we can observe a great variability when changing this quantile order.

Table 2: *Mean Squared Error of extreme quantile-based regularized estimators over 2000 Monte-Carlo simulations for the Uniform case, true values are* $\varphi_0 = 1$, $\rho_0 = 2$ *and* $\ell_0 = 1$. *From Daouia et al. (2010) and the reference [3] therein.*

Sample size	$\hat{\rho}$	α_n^*	$\tilde{\varphi}_{\alpha_n^*}$	α_n^*	$\tilde{\varphi}_{\alpha_n^*}(\rho_0)$	α_n^*
$N_x = 62$	836.96814	0.858	3.45696	0.889	0.00135	0.868
$N_x = 250$	118.17269	0.767	3.84698	0.767	0.00139	0.812
$N_x = 562$	1.28492	0.767	0.08546	0.782	0.00138	0.767
$N_x = 1000$	0.65257	0.782	0.07620	0.798	0.00140	0.830
$N_x = 5000$	0.04085	0.750	0.00552	0.812	0.00028	0.837

We observe a gain of the MSE when estimating ρ by a factor of the order 1000, 8, 6 and 1.3 for the samples sizes 100, 500, 1000 and 5000, respectively. Remember that we selected in Table 2 the best order of the quantile-type estimator, whereas we selected here the order $m = N_x^{1/3}$ and $a = 10$ given by our simple rule. Of course, better results could be obtained, case by case, for other choices of m and a.

For the estimation of the frontier point, the gain of the MSE has a factor of the order 450, 70, 150 and 60 (respectively). We also observe some qualitative gain for the estimation of

the frontier when ρ is known, but here the gain is of a factor ranging from 2.5 to 5 when N_x goes from 500 to 5000. Again in this comparison, we selected the best value of the quantile order in the results from the Tables in Daouia et al.

To conclude this general comparison between the two approaches (using the order-m functions here and using the order-α quantiles as in Daouia et al., 2010), we can say that we have much better results in these particular simulations and that with the approach here, we gain a lot in terms of the stability of the estimators with respect to the choice of the regularization parameters. Tables 1 and 3 in Daouia et al. indicate indeed a huge sensitivity to the choice of the quantile order when defining the base estimator (the MSE can change by a factor of several thousands if the wrong order α_n is picked out) and this is not the case here where we observe a great stability in the estimation of the frontier.

5.1.2 Beta densities for the efficiency term

Now, we analyze the results with different behaviors of the density of the efficiencies at the frontier points (density tending to infinity, having a jump or converging to zero at the frontier points). We select the following model $Y = X\,V$ where $X \sim \text{Unif}(0,1)$ and $V \sim \text{Beta}(\beta,\beta)$ with values of $\beta = 0.5, 1$ and 3. Note that in all the cases, $\mathbb{E}(V) = 0.5$. Again we focus the results for the value $x = 1$, so that $N_x = n$. The results are shown in Tables 3 to 5. In the first case the density tends to infinity at the frontier, and the FDH estimator $\hat{\varphi}$ should be performant. It is indeed the case, but our regularized estimator $\tilde{\varphi}$ performs even slightly better for $N_x = 100$ and much better for larger N_x reaching less Bias and MSE (when choosing $a = 10$). Again, the estimation of ρ and ℓ is more difficult but our simple rule of thumb ($m = N_x^{1/3}$ and $a = 10$) shows nice behavior of the estimators. When β increases (jump at the frontier for $\beta = 1$ and going smoothly to zero when $\beta = 3$), the results for the estimators of the frontier deteriorate a little. However, as expected, our regularized estimator $\tilde{\varphi}$ remains still better (with $a = 10$) than the FDH estimator $\hat{\varphi}$ and the CFS unregularized order-m_n frontier estimator $\tilde{\varphi}_{CFS}$, in terms of both bias and MSE.

As for the uniform case above, when $a = 2$ we have less remarkable results for the two first cases ($\beta = 0.5, 1$), but when the density is going smoothly to zero at the frontier ($\beta = 3$), the value $a = 2$ provides better results as far as N_x is larger than 500. This indicates again how difficult is the problem of selecting an optimal a. In this latter case, we illustrate the estimation of the frontier in the full range of X in the next section, and we analyze the robustness properties of our estimator and investigate the achieved coverage of the normal confidence intervals suggested above.

Table 3: *Bias and Mean Squared Error (MSE) of the estimates over 1000 Monte-Carlo simulations: case of the Beta*$(0.5, 0.5)$*, true values are* $\varphi_0 = 1$*,* $\rho_0 = 1.5$ *and* $\ell_0 = 0.4244$*.*

	$N_x = 100$		$N_x = 500$		$N_x = 1000$		$N_x = 5000$	
$a = 2$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	1.198985	3.023330	0.650579	0.548628	0.489587	0.288141	0.291050	0.094974
$\hat{\ell}$	0.330865	0.192974	0.315225	0.108012	0.295414	0.090897	0.225762	0.052464
$\hat{\ell}(\rho_0)$	0.482618	0.252249	0.278188	0.080796	0.222635	0.051233	0.133310	0.018186
$\tilde{\varphi}$	0.121722	0.058908	0.050483	0.005139	0.032115	0.001843	0.014537	0.000314
$\tilde{\varphi}(\rho_0)$	-0.090905	0.011651	-0.048410	0.002867	-0.035330	0.001506	-0.016387	0.000310
$a = 10$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	-0.063176	0.168646	0.146422	0.111090	0.123820	0.061474	0.096074	0.021250
$\hat{\ell}$	0.310572	0.124454	0.237085	0.075298	0.201660	0.052083	0.141359	0.024371
$\hat{\ell}(\rho_0)$	0.341740	0.138351	0.188876	0.039628	0.153513	0.025683	0.092683	0.009117
$\tilde{\varphi}$	-0.053538	0.005929	-0.007142	0.000553	-0.004477	0.000217	-0.000053	0.000022
$\tilde{\varphi}(\rho_0)$	-0.048796	0.005060	-0.020620	0.000787	-0.014593	0.000391	-0.005882	0.000062
$\tilde{\varphi}_{CFS}$	-0.096222	0.011416	-0.033583	0.001368	-0.022219	0.000602	-0.007281	0.000065
$\hat{\varphi}$	-0.072121	0.007519	-0.024156	0.000848	-0.016576	0.000391	-0.005347	0.000042

Table 4: *Bias and Mean Squared Error (MSE) of the estimates over 1000 Monte-Carlo simulations: case of the Beta*$(1, 1)$*, true values are* $\varphi_0 = 1$*,* $\rho_0 = 2$ *and* $\ell_0 = 0.5$*.*

	$N_x = 100$		$N_x = 500$		$N_x = 1000$		$N_x = 5000$	
$a = 2$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	1.735601	9.218976	0.886403	1.155148	0.682494	0.628021	0.417289	0.208678
$\hat{\ell}$	0.345564	0.363425	0.335707	0.146882	0.336649	0.125037	0.291430	0.087952
$\hat{\ell}(\rho_0)$	0.665852	0.517782	0.390621	0.163892	0.315638	0.105073	0.200225	0.041405
$\tilde{\varphi}$	0.183781	0.188113	0.072240	0.013123	0.049637	0.005367	0.024118	0.000999
$\tilde{\varphi}(\rho_0)$	-0.091720	0.012735	-0.052870	0.003677	-0.039797	0.001992	-0.020937	0.000525
$a = 10$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	-0.331901	0.444470	0.086348	0.248021	0.109707	0.135392	0.138556	0.065002
$\hat{\ell}$	0.359753	0.193566	0.280033	0.109485	0.253591	0.087861	0.201901	0.053630
$\hat{\ell}(\rho_0)$	0.506769	0.342331	0.273330	0.088737	0.222238	0.055914	0.141044	0.021551
$\tilde{\varphi}$	-0.091613	0.015110	-0.018562	0.002537	-0.009895	0.001006	0.000817	0.000188
$\tilde{\varphi}(\rho_0)$	-0.059530	0.007942	-0.027354	0.001641	-0.019782	0.000782	-0.009198	0.000161
$\tilde{\varphi}_{CFS}$	-0.150452	0.025881	-0.069210	0.005540	-0.049630	0.002802	-0.021428	0.000531
$\hat{\varphi}$	-0.120797	0.018561	-0.055114	0.003931	-0.039773	0.001989	-0.017015	0.000375

Table 5: *Bias and Mean Squared Error (MSE) of the estimates over 1000 Monte-Carlo simulations: case of the Beta(3, 3), true values are* $\varphi_0 = 1$, $\rho_0 = 4$ *and* $\ell_0 = 2.5$.

	$N_x = 100$		$N_x = 500$		$N_x = 1000$		$N_x = 5000$	
$a = 2$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	0.671507	13.664936	-0.021354	1.636235	-0.130470	0.875667	-0.243409	0.295094
$\hat{\ell}$	-0.761246	1.889056	-0.920891	1.177173	-0.912659	1.022306	-0.910911	0.863933
$\hat{\ell}(\rho_0)$	-0.389781	1.442638	-0.787752	0.832138	-0.829105	0.817469	-0.858793	0.775039
$\tilde{\varphi}$	0.090425	0.221473	0.014391	0.022011	0.004275	0.011064	-0.004175	0.002342
$\tilde{\varphi}(\rho_0)$	-0.000367	0.006222	0.015008	0.001761	0.016361	0.001176	0.017433	0.000518
$a = 10$	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
$\hat{\rho}$	-2.053562	4.921744	-0.825120	4.708721	-0.630922	2.096024	-0.419817	0.647445
$\hat{\ell}$	-1.068665	1.595360	-1.057668	1.413049	-1.042189	1.310455	-0.967238	1.047182
$\hat{\ell}(\rho_0)$	0.348089	5.674396	-0.585639	0.849339	-0.670422	0.725935	-0.742065	0.634133
$\tilde{\varphi}$	-0.199484	0.051431	-0.063119	0.030965	-0.041521	0.013676	-0.018208	0.003018
$\tilde{\varphi}(\rho_0)$	-0.019372	0.008809	0.005711	0.002002	0.008604	0.001276	0.011473	0.000432
$\bar{\varphi}_{CFS}$	-0.267260	0.074765	-0.175616	0.032211	-0.147603	0.022763	-0.097637	0.009952
$\hat{\varphi}$	-0.237955	0.061255	-0.155909	0.026241	-0.131248	0.018605	-0.086933	0.008159

5.2 Estimation of the frontier function

5.2.1 One simulated sample

We first illustrate the behavior of the frontier estimate in the case of a beta density for the efficiencies, with the model described in the preceding subsection. We show the case where the density is converging smoothly to zero at the frontier ($\beta = 3$). For estimating the frontier function over the full range of X, it is common to assume that the tail index $\rho_x = \rho$ is constant[6] in x over the range of X (which is true in the simulated scenario). We estimate this value by a weighted mean (the weights are N_x) of the local values $\hat{\rho}_x$ computed over a fixed grid of 10 values of x from 0.25 till 1. In this step of estimating ρ_x from a simulated sample of size $n = 1000$, we keep $m = N_x^{1/3}$ but we choose $a = 2$ (due to the good results in the Monte-Carlo experiments above when $\beta = 3$). We obtained the value 4.1553 where the true value is 4. Then we compute the values $\hat{\ell}_x$ and $\widetilde{\varphi}_m(x)$ on a grid of 76 values for x from 0.25 till 1. The 95% confidence intervals for each value of x were obtained by using the normal approximation, centered on $\widetilde{\varphi}_m(x)$ and the variance is estimated by a bootstrap algorithm (200 replications: in each bootstrap sample, the tail index is re-estimated over the same grid of 10 values of x and averaged to mimic the original estimation procedure). We compute also the CFS unregularized order-m estimator $\widehat{\varphi}_{m_n}(x)$ (where we set as above $m_n = N_x$) with the Weibull pointwise confidence intervals for the frontier $\varphi(x) = x$. The results are quite sensible and are displayed on the top panel of Figure 1. We see that our estimate is

[6]The shape parameter $\beta_x = \rho_x - (p + 1)$ of the joint density of (X, Y) (see Corollary 2.2 in Daouia et al. (2010)) is often assumed to be either known and independent of x in parametric approaches or null in nonparametric approaches. It is more reasonable and less restrictive to keep only the condition of independance from x and then to estimate $\beta_x = \beta$ or equivalently $\rho_x = \rho$.

better than the FDH estimate and the CFS estimate (closer to the true frontier). We see clearly that the pointwise confidence intervals cover the true frontier for both our regularized "normal" confidence intervals and for the unregularized "Weibull" case. The latter however seem to be narrrower (achieved coverages will be estimated below). It appears in this sample that the FDH estimator is even outside the 95% confidence intervals for all x.

In the two next panels of Figure 1, in order to investigate resistance to outliers, we introduce in the middle panel one important outlier at $(x, y) = (0.5, 0.8)$ (represented by a black circle) and in the bottom panel, 3 outliers at input values $x = (0.25, 0.50, 0.75)$ having output 20% above the frontier levels, so $y = (0.30, 0.60, 0.90)$. These outliers where introduced only for estimating the frontier levels (keeping our original estimate of ρ).

We can appreciate the robustness of our frontier estimates and their corresponding confidence intervals (relative to the FDH estimator and even to the CFS estimator with its Weibull confidence intervals). Of course, in practice, we could easily detect this outlier (even for dimension $p > 1$, because it is far outside the confidence interval obtained from our robust regularized estimator at this point). Once this is observed, and as always when detecting potential outliers, this point could be removed from the sample only after a careful analysis. But it is remarkable how the regularized estimator is resistant to the different added outliers. The behavior of the Weibull confidence intervals is, however, disappointing in terms of robustness since both confidence bands miss the target for many values of x. This will be confirmed in the Monte-Carlo experiment below that will estimate the coverages of the resulting confidence intervals.

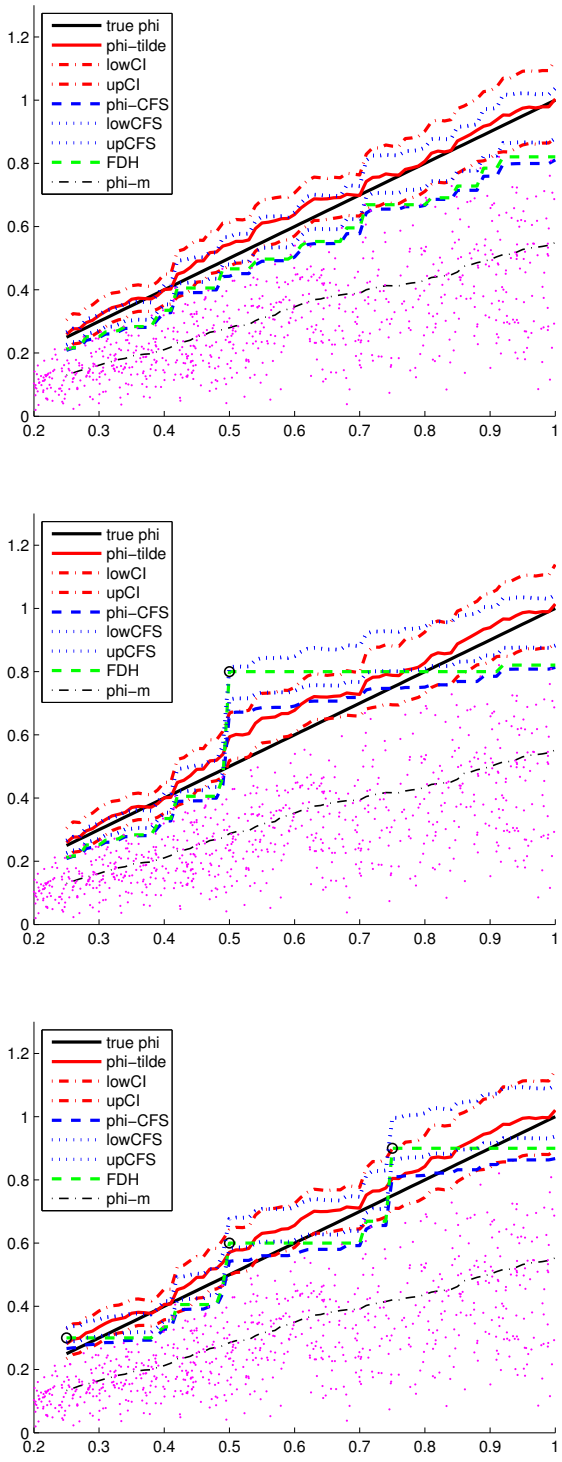

Figure 1: *Linear frontier:* $Y = X\,V$ *with* $V \sim \text{Beta}(3,3)$ *and* $n = 1000$. *Middel panel, with one outlier and bottom panel with 3 outliers. The base (biased) estimator 'phi-m' is* $\hat{\varphi}_{m_n}(x)$, *'phi-CFS' is the unregularized estimator* $\tilde{\varphi}_{CFS}$ *and our regularized estimator 'phi-tilde' is* $\tilde{\varphi}(x)$.

5.2.2 Coverage of Confidence Intervals

Here we investigate whether the procedure for estimating the confidence intervals for $\varphi(x)$ based on the bias-corrected regularized estimator (using normal tables) provides estimates with reasonable coverages. In the same time, we will compare the performance of this procedure with the confidence intervals obtained from the unregularized order-m frontier estimator which follows the Weibull distribution. We will present the same scenarios of the preceding example where $V \sim Beta(3,3)$, without and with outliers. We focus the analyis at 4 equidistant values of $x = (0.25, 0.50, 0.75, 1.00)$, and provide also in Table 6 the Bias and MSE of the two estimators. Here, as above, $m = N_x^{1/3}$ and $a = 2$.

Consider first the case where there are no outliers. We see that the $Bias$ and MSE for both $\widetilde{\varphi}_m(x)$ and $\tilde{\varphi}_{CFS}(x)$ increase with x: this is due to our Monte-Carlo setup with the multiplicative model $Y = XV$, there is much less variation in Y when X is small. But we see that our regularized estimator (using the first step estimate of ρ) has better behavior in terms of bias and MSE. For the coverages, we see that the estimator of the normal confidence intervals derived from our regularized estimator has a good coverage for all x and much better than the estimated confidence intervals obtained from the Weibull (remember that here too, we need to estimate ρ and ℓ_x, we used the same values as the ones used for our regularized estimate).

A better resistance to outliers of our regularized estimator appears clearly in the two bottom blocks of Table 6 when comparing the achieved coverages. Even if the robustness of the unregularized estimator is somewhat preserved (the bias and MSE are not so bad and much better than the FDH estimator whose MSE, not reproduced here to save place, are multiplied by a factor 10 near the outlying points), the obtained confidence intervals miss completely the target near the values of x where the outliers are introduced. This already appeared in Figure 1 for one particular sample of size $n = 1000$.

Table 6: *Comparison of regularized Normal estimator (using $\tilde{\varphi}_m(x)$) and unregularized Weibull estimator (using $\tilde{\varphi}_{CFS}(x)$). Coverages (cov), Average Lengths (avl) of 95% Estimated Confidence Intervals, Bias and Mean Squared Error (MSE) of frontier estimates at 4 values of x. Results obtained over 1000 Monte-Carlo simulations with $n = 1000$. Case where $V \sim Beta(3,3)$ ($\rho_0 = 4$) with $\varphi_0(x) = x$.*

x	$cov_{\widetilde{\varphi}_m}$	$avl_{\widetilde{\varphi}_m}$	$cov_{\tilde{\varphi}_{CFS}}$	$avl_{\tilde{\varphi}_{CFS}}$	$Bias_{\widetilde{\varphi}_m}$	$MSE_{\widetilde{\varphi}_m}$	$Bias_{\tilde{\varphi}_{CFS}}$	$MSE_{\tilde{\varphi}_{CFS}}$
No outliers								
0.25	0.9590	0.0741	0.5180	0.0379	0.0015	0.0003	-0.0529	0.0029
0.50	0.9690	0.1265	0.7110	0.0765	0.0056	0.0009	-0.0883	0.0082
0.75	0.9810	0.1659	0.8570	0.1154	0.0107	0.0013	-0.1181	0.0146
1.00	0.9690	0.2488	0.8640	0.1542	0.0151	0.0031	-0.1475	0.0227
One outliers at $x = 0.50$, $y = 0.80$								
0.25	0.9590	0.0741	0.5180	0.0379	0.0015	0.0003	-0.0529	0.0029
0.50	0.7200	0.1399	0	0.0909	0.0585	0.0044	0.1573	0.0248
0.75	0.9720	0.1718	0.0080	0.1222	0.0356	0.0025	-0.0118	0.0002
1.00	0.9730	0.2496	0.8780	0.1562	0.0232	0.0034	-0.1462	0.0223
Three outliers at $x = 0.25, 0.50, 0.75$, $y = 0.30, 0.60, 0.90$								
0.25	0.8590	0.0782	0	0.0438	0.0251	0.0009	0.0122	0.0002
0.50	0.9350	0.1322	0	0.0842	0.0356	0.0021	0.0308	0.0010
0.75	0.9170	0.1754	0	0.1262	0.0513	0.0039	0.0522	0.0028
1.00	0.9790	0.2512	0.9460	0.1577	0.0296	0.0038	-0.1158	0.0136

5.2.3 A real data example

We use the same real data set as in Cazals et al. (2002) and Daouia et al. (2010) on the frontier analysis of 9521 French post offices observed in 1994, with X as the quantity of labor and Y as the volume of delivered mail. In this illustration, we only consider the $n = 4000$ observed post offices with the smallest levels x_i.

We first start by assuming, as in most nonparametric econometric frontier studies, that the joint density of (X, Y) has a jump on the frontier, so $\rho_x = p + 1 = 2$. The cloud of points and the resulting estimates are provided in Figure 2. The FDH estimator is clearly determined by only a few very extreme points. If we delete the 4 anomalous observations (represented by circles in the figure) from the sample, we obtain the picture of the right panel: the FDH estimator changes drastically, whereas the regularized estimator, with $m_n = N_x^{1/3}$ and $a = 2$, is very robust to the presence of these 4 extreme points. Although its nice behavior, the unregularized estimator is less resistant to the extreme points, (here we set, as in the simulations, $m_n = N_x$). Again the confidence intervals obtained by using the regularized estimator $\widetilde{\varphi}_m(x)$ were obtained by a bootstrap algorithm. We observe very narrow confidence intervals for $\varphi(x)$, this is due to the fact that ρ is fixed. Also, looking to these two pictures, it seems that $\rho = 2$ is a too strong assumption: the regularized estimator is very far from the border of the cloud of points. This will be corrected below by estimating ρ, but more noise will be caused.

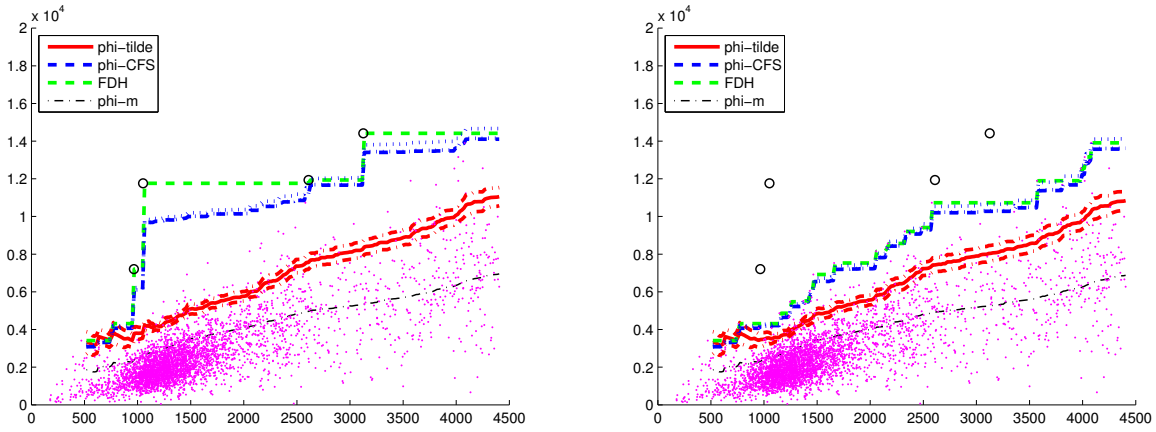

Figure 2: *Resulting estimator $\widetilde{\varphi}_m(x)$ for the French post offices with $\rho = 2$. In the left panel, the 4 anomalous points (circles) are used in the estimation of the two frontier functions.*

For estimating ρ_x, we proceed as above by assuming that $\rho_x = \rho$ is an unknown constant and we average the values of $\hat{\rho}_x$ obtained over a grid of 20 values of x (again with a weighted mean). Here, larger values of m were needed for computing $\hat{\rho}_x$, to avoid numerical instabilities in (4.1): we choose $m = 10N_x^{1/3}$ and we set $a = 8$. In this first step estimation of ρ, we only used the sample without the 4 outliers detected above. This provided the estimator $\hat{\rho} = 3.5573$, indicating that the density of the efficiencies is tending to zero at the frontier, but not its first derivative; a reasonable result when looking to the cloud of data points in Figure 2.

Then, for estimating the frontier, we proceed as usual with the full sample, keeping the basic rule of thumb $m = N_x^{1/3}$ and $a = 2$, as in the Monte-Carlo exercices above. The results are displayed on the left panel of Figure 3, where we see that the higher value of $\hat{\rho}$ (compared to $\rho = 2$ in Figure 2) pushes our estimator to the North, as expected, because the correction for the bias is larger. We also observe that the 4 outliers are left outside our 95% upper confidence band and that the confidence intervals obtained via the unregularized CFS estimator (Weibull case) are again really outside the observed cloud of points except for these 4 extreme points.

The right panel of the figure, where the 4 extremes are excluded from the sample, indicates how the frontier estimate is robust to the outliers (as compared to FDH and CFS). We observe that most of the FDH frontier and of the CFS unregularized frontier is now inside the 95% confidence intervals, when the 4 outliers have been dropped out of the sample.

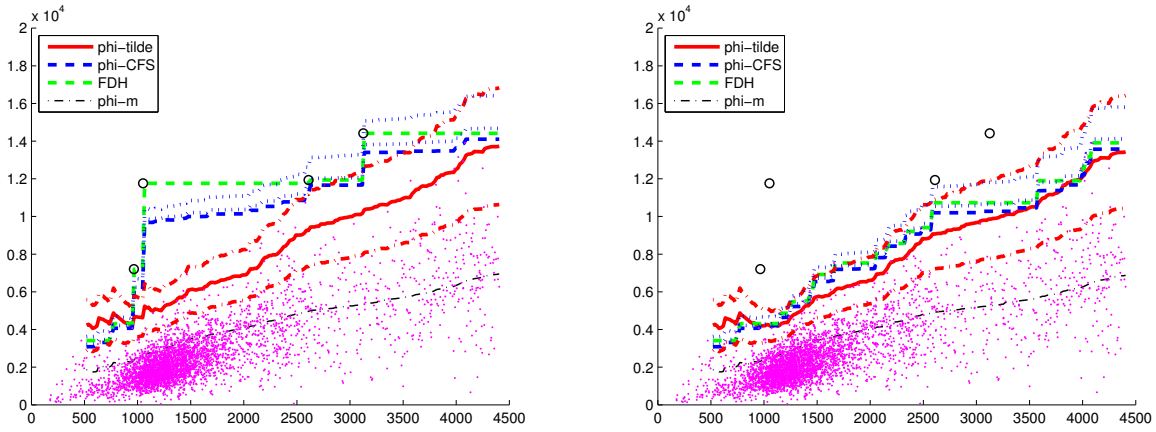

Figure 3: *Resulting estimator* $\widetilde{\varphi}_m(x)$ *for the postal data with* $\hat{\rho} = 3.5573$. *In the left panel, the 4 anomalous points (circles) are used in the estimation of the two frontier functions.*

6 Conclusions

We have derived in this paper the theory of an estimator of a frontier function having an asymptotic normal distribution. The basic tool is the order-m partial frontier where we let the order m to converge to infinity when $n \to \infty$ but at a slow rate. Indeed, if the rate is too fast, the order-m frontier will converge too quickly to the full frontier and the corresponding estimator will converge to the FDH estimator, having a Weibull limiting distribution. The final estimator is then corrected for its inherent bias. We thus can view our estimator as a regularized frontier estimator which, in addition, is more robust to extreme values and outliers than the usual nonparametric frontier estimators, including the unregularized order-m estimator of Cazals et al. (2002) converging toward a Weibull distribution.

In addition, if the tail index ρ_y and the behavior of the conditional distribution of X given that $Y \geq y$ near the frontier points is not known (ℓ_y), we provide an easy way to estimate them consistently.

The performances of our estimators are evaluated in finite samples through some Monte-Carlo experiments, showing very nice regular behavior of the estimators, in particular for the estimator of the frontier. We also illustrate how to provide, in an easy way, confidence intervals for the frontier function in a simulated data set where the FDH estimator gives very poor results. Some Monte-Carlo experiments indicate reasonable coverages of the resulting confidence intervals. We also illustrate our procedure with a real data set from the French Post Offices.

Important research issues are still open and deserve for future work. This includes a way

for selecting optimal regularization parameters m and a, which is particularly important for deriving the estimator of the tail index ρ_y. But this is known as a hard mathematical problem in extreme-value theory. Once ρ_y is well estimated (or assumed to be known), the estimate of the frontier itself is much more robust to the choice of the order m. Another trail of research would be to define estimators of ρ_y, ℓ_y and $\varphi(y)$ when they are considered as smoothed functions of y.

Appendix: The Output Oriented Case

In this section we only give the useful notations and formulas for the output oriented case. Here the attainable production set is defined as $\Psi = \{(x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+ \mid x \text{ can produce } y\}$ and the production frontier is represented by the graph of the production function $\varphi(x) = \sup\{y \,|\, (x, y) \in \Psi\}$. The distribution function of (X, Y) can be denoted $F(x, y)$ and $F(\cdot|x) = F(x, \cdot)/F_X(x)$ will be used to denote the conditional distribution function of Y given $X \leq x$, with $F_X(x) = F(x, \infty) > 0$. It has been proven in Cazals et al. (2002) that under the free disposability assumption, the production function can equivalently be defined by

$$\varphi(x) = \sup\{y \geq 0 | F(y|x) < 1\} \tag{A.1}$$

The order-m partial frontier is now defined as

$$\varphi_m(x) = \mathbb{E}\big[\max(Y_1, \ldots, Y_m)|X \leq x\big], \tag{A.2}$$

where $(Y_1, \ldots, Y_m)$ are m i.i.d. random variables generated by the conditional distribution of Y given $X \leq x$. It is shown in Cazals et al. that $\varphi_m(x) = \int_0^\infty \big(1 - [F(u|x)]^m\big)\, du = \varphi(x) - \int_0^{\varphi(x)} [F(u|x)]^m\, du$, so that $\varphi_m(x) \to \varphi(x)$ as $m \to \infty$.

Nonparametric estimators of these frontiers are obtained by plugging the empirical version of the unknown distribution $F(\cdot|x)$ in the definition above. So we obtain

$$\hat{\varphi}(x) = \sup\{y \geq 0 | \widehat{F}(y|x) < 1\} = \max_{\{i:X_i \leq x\}} Y_i \tag{A.3}$$

$$\hat{\varphi}_m(x) = \hat{\varphi}(x) - \int_0^{\hat{\varphi}(x)} [\widehat{F}(u|x)]^m\, du, \tag{A.4}$$

where $\widehat{F}(y|x) = \widehat{F}(x, y)/\widehat{F}_X(x)$ with $\widehat{F}(x, y) = 1/n \sum_{i=1}^n \mathbb{1}(X_i \leq x, Y_i \leq y)$ and $\widehat{F}_X(x) = 1/n \sum_{i=1}^n \mathbb{1}(X_i \leq x)$. For any given x and a fixed value of m, we have as $n \to \infty$,

$$\frac{\sqrt{n}}{\sigma(m, x)}\big(\hat{\varphi}_m(x) - \varphi_m(x)\big) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \tag{A.5}$$

where the variance can be written, as in (3.2), as

$$\sigma^2(m,x) = \frac{2m^2}{F_X(x)} \int_0^{\varphi(y)} \int_0^{\varphi(y)} F^m(y|x) F^{m-1}(u|y)(1-F(u|x))\, \mathbb{1}(y \leq u)\, dy\, du. \qquad (A.6)$$

The regularity condition can be written here as

$$F_X(x)(1-F(y|x)) = \ell_x(\varphi(x)-y)^{\rho_x} + o(\varphi(x)-y)^{\rho_x}, \quad \text{as } y \uparrow \varphi(x), \qquad (A.7)$$

where $\ell_x > 0$, $\rho_x > p$ and $\varphi(x)$ is differentiable in x with strictly positive first partial derivatives. Then, from the equation (2.5) in Daouia et al. (2010), we obtain the useful relation, as $m \to \infty$,

$$\varphi(x) - \varphi_m(x) = \Gamma\left(1+\frac{1}{\rho_x}\right)\left(\frac{1}{m\,\ell_x}\right)^{1/\rho_x} + o(m^{-1/\rho_x}). \qquad (A.8)$$

Then, the asymptotic theory given in Sections 3 and 4 can be easily adapted.

References

[1] Aragon, Y., A. Daouia and C. Thomas-Agnan (2005), Nonparametric Frontier Estimation: A Conditional Quantile-based Approach, *Econometric Theory*, 21, 358–389.

[2] Cazals, C., J.P. Florens and L. Simar (2002), Nonparametric frontier estimation: a robust approach, *Journal of Econometrics*, 106, 1–25.

[3] Charnes, A., Cooper W.W. and E. Rhodes (1978), Measuring the inefficiency of decision making units, *European Journal of Operational Research* 2 (6), 429–444.

[4] Daouia, A., J.P. Florens and L. Simar (2010), Frontier estimation and Extreme value theory, *Bernoulli*, 16, 1039–1063.

[5] Daouia, A. and I. Gijbels (2011a), Robustness and inference in nonparametric partial-frontier modeling, *Journal of Econometrics*, 161, 147–165.

[6] Daouia, A. and I. Gijbels (2011b), Estimating frontier cost models using extremiles. In : *Exploring research frontiers in contemporary statistics and econometrics - Festschrift in honor of L. Simar*, ed. by P.W. Wilson and I. Van Keilegom, Springer.

[7] Daouia, A. and L. Simar (2005), Robust Nonparametric Estimators of Monotone Boundaries, *Journal of Multivariate Analysis*, 96, 311–331.

[8] Daouia, A. and L. Simar (2007), Nonparametric efficiency analysis: a multivariate conditional quantile approach, *Journal of Econometrics*, 140, 375–400.

[9] Daraio, C. and L. Simar (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*, Springer, New-York.

[10] Deprins, D., Simar, L. and H. Tulkens (1984), Measuring labor inefficiency in post offices. In *The Performance of Public Enterprises: Concepts and measurements*. M. Marchand, P. Pestieau and H. Tulkens (eds.), Amsterdam, North-Holland, 243–267.

[11] Farrell, M.J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society*, A(120), 253–281.

[12] Ferreira, A., de Haan, L. and L. Peng (2003). On optimizing the estimation of high quantiles of a probability distribution. *Statistics*, 37, 401–434.

[13] Kneip, A., B. Park and L. Simar (1998), A Note on the Convergence of Nonparametric DEA Estimators for Production Efficiency Scores, *Econometric Theory*, 14, 783–793.

[14] Kneip, A, L. Simar and P.W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory*, 24, 1663–1697.

[15] Park, B. Simar, L. and Ch. Weiner (2000), The FDH Estimator for Productivity Efficiency Scores : Asymptotic Properties, *Econometric Theory*, Vol 16, 855–877.

[16] Shephard, R.W. (1970). *Theory of Cost and Production Function*. Princeton University Press, Princeton, New-Jersey.

[17] Simar, L. and P.W. Wilson (2008), Statistical Inference in Nonparametric Frontier Models: recent Developments and Perspectives, in *The Measurement of Productive Efficiency*, 2nd Edition, Harold Fried, C.A.Knox Lovell and Shelton Schmidt, editors, Oxford University Press.